

Inv: SoBEK et al.

CRFE

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

R Date: 3/9/01

SEARCH REQUEST FORM

113445

Requestor's Name: JANE ZARA

Serial Number: 09/803,165

Date: 2-6-04

Phone: 2-0765

Art Unit: 1635

2018

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

11/81

Please Search Seq ID No: 34

No size limits

Thanks

STAFF USE ONLY

Date completed: 2/10
Searcher: _____
Terminal time: _____
Elapsed time: _____
CPU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Search Site

____ STIC
____ CM-1
____ Pre-S

Type of Search

____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors

____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 11:48:25 ; Search time 28 Seconds
(without alignments)
1169.593 Million cell updates/sec

Title: US-09-803-165-34
Perfect score: 4026
Sequence: 1 MIFDIDYITKDGKPIIRFK.....KEDLKYSKQVCLDAWLKK 774

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3225.5	80.1	774	3	US-08-902-632-2
2	3222.5	80.0	774	3	US-09-073-354-1
3	3222.5	80.0	774	3	US-08-656-005A-1
4	3222.5	80.0	774	3	US-09-073-259-1
5	3222.5	80.0	774	3	US-09-363-095-1
6	3222.5	80.0	774	3	US-09-418-027-1
7	3209.5	79.7	778	2	US-08-906-925-4
8	3151	78.3	775	1	US-07-966-278-1
9	3151	78.3	775	1	US-08-424-921-1
10	3151	78.3	775	2	US-08-556-355A-1
11	3151	78.3	775	2	US-07-803-627A-1
12	3151	78.3	775	4	US-09-244-889A-1
13	3151	78.3	776	2	US-08-688-649-37
14	3141.5	78.0	773	4	US-09-585-858-35
15	3139	78.0	779	1	US-08-375-134-12
16	3139	78.0	779	5	PCT-US95-15263-12
17	2974	73.9	727	4	US-09-585-858-32
18	2502.5	62.2	1022	2	US-08-271-364A-8
19	2502.5	62.2	1022	1	US-08-222-715B-27
20	2069	51.4	1019	1	US-08-271-364A-7
21	2069	51.4	1019	2	US-08-222-715B-26
22	1595	39.6	396	1	US-08-229-284A-2
23	1403	34.8	788	2	US-08-907-166-6
24	1403	34.8	788	4	US-09-391-340-6
25	1194	29.7	803	2	US-08-907-166-4
26	1194	29.7	803	4	US-09-391-340-4
27	1191	29.6	803	1	US-08-062-368-4

28	1180	29.3	803	1	US-08-062-368-2
29	656	16.3	734	4	US-09-585-858-8
30	611	15.2	738	4	US-09-585-858-7
31	611	15.2	1107	1	US-08-366-577-2
32	611	15.2	1107	5	PCT-US96-00005-2
33	602	15.0	735	4	US-09-585-858-10
34	602	15.0	738	4	US-09-585-858-6
35	599.5	14.9	734	4	US-09-585-858-9
36	591	14.7	877	2	US-08-907-166-8
37	591	14.7	877	4	US-09-391-340-8
38	588.5	14.6	762	2	US-08-907-166-10
39	588.5	14.6	762	4	US-09-391-340-10
40	585.5	14.5	702	4	US-09-585-858-33
41	585.5	14.5	1462	3	US-07-792-600-31
42	585.5	14.5	1462	3	US-09-157-021-31
43	585.5	14.5	1462	3	US-09-156-842-31
44	582	14.5	757	4	US-09-585-858-23
45	568.5	14.1	757	4	US-09-585-858-24

ALIGNMENTS

RESULT 1
US-08-902-632-2
; Sequence 2, Application US/08902632
; Patent No. 6008025
; GENERAL INFORMATION:
; APPLICANT: KOMATSUBARA, Hideyuki
; APPLICANT: KITABAYASHI, Masao
; APPLICANT: KAMIMURA, Hideki
; APPLICANT: KAWAKAMI, Bunsei
; APPLICANT: KAWAMURA, Yoshihisa
; APPLICANT: TAKAGI, Masahiro
; APPLICANT: IMANAKA, Tadayuki
; TITLE OF INVENTION: Modified Thermostable DNA Polymerase,
; TITLE OF INVENTION: and DNA Polymerase Composition for Nucleic Acid
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1 Broadway
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 inch floppy disk, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1 Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,632
; FILING DATE: Concurrent Herewith
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 198911/96
; FILING DATE: 29-JUL-1996
; APPLICATION NUMBER: JP 200446/96
; APPLICATION NUMBER: 30-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: GREASON, Edward W.
; REGISTRATION NUMBER: 18,918
; REFERENCE/DOCKET NUMBER: 2418/7
; TELEPHONE: 212-425-7200
; TELEFAX: 212-425-5288
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-902-632-2		Query Match		80.1%; Score 3225.5; DB 3; Length 774;			
		Best Local Similarity		77.5%; Pred. No. 1.2e-242;			
		Matches 599; Conservative		82; Mismatches 89; Indels 3; Gaps 2;			
		TITLE OF INVENTION: Amplifying Nucleic Acids					
		NUMBER OF SEQUENCES: 16					
		CORRESPONDENCE ADDRESS:					
		ADDRESSEE: Kenyon & Kenyon					
		STREET: 1025 Connecticut Avenue, N.W., Suite 600					
		CITY: Washington					
		STATE: D.C.					
		COUNTRY: USA					
		ZIP: 20036					
		COMPUTER READABLE FORM:					
		MEDIUM TYPE: 3.5" Floppy disk					
		COMPUTER: IBM PC compatible					
		OPERATING SYSTEM: PC-DOS/MS-DOS					
		SOFTWARE: WordPerfect 6.1 Windows					
		CURRENT APPLICATION DATA:					
		APPLICATION NUMBER: US/09/073,354					
		FILING DATE: Concurrent Herewith					
		CLASSIFICATION: 0506					
		PRIOR APPLICATION DATA:					
		APPLICATION NUMBER: 08/656,005					
		FILING DATE: 24 MAY 1996					
		APPLICATION NUMBER: JP 134096/95					
		FILING DATE: 31 MAY 1995					
		ATTORNEY/AGENT INFORMATION:					
		NAME: Toffenetti, Judith L.					
		REGISTRATION NUMBER: 39,048					
		REFERENCE/DOCKET NUMBER: 2418/9					
		TELECOMMUNICATION INFORMATION:					
		TELEPHONE: 202-429-1776					
		TELEFAX: 202-429-0796					
		INFORMATION FOR SEQ ID NO: 1:					
		SEQUENCE CHARACTERISTICS:					
		LENGTH: 774 amino acids					
		TYPE: amino acid					
		STRANDEDNESS: double					
		TOPOLOGY: linear					
		MOLECULE TYPE: protein					
		US-09-073-354-1					
		Query Match		80.0%; Score 3222.5; DB 3; Length 774;			
		Best Local Similarity		77.4%; Pred. No. 2e-242;			
		Matches 599; Conservative		83; Mismatches 89; Indels 3; Gaps 2;			
Qy	1	MIFDTDYITKDGKPIIRIFKKNKGEFKIELDPHPQPIYVALLKDDSAIDBIKAIKGERHG	60	Qy	1	MIFDTDYITKDGKPIIRIFKKNKGEFKIELDPHPQPIYVALLKDDSAIDBIKAIKGERHG	60
Db	1	MILDTDYITDGKPIIRIFKKNKGEFKIELDRTFEPFYALLKDDSAIEVKKITAERHG	60	Db	1	MILDTDYITDGKPIIRIFKKNKGEFKIELDRTFEPFYALLKDDSAIEVKKITAERHG	60
Qy	61	KIVRVDAVKKKKFLGRDVEVWKLIFHPQDVPAIRKIREHPAVIDIYEDIDIPAKRY	120	Qy	61	KIVRVDAVKKKKFLGRDVEVWKLIFHPQDVPAIRKIREHPAVIDIYEDIDIPAKRY	120
Db	61	TVVTVKRVKVKQKFLGRDVEVWKLIFHPQDVPAIRKIREHPAVIDIYEDIDIPAKRY	120	Db	61	TVVTVKRVKVKQKFLGRDVEVWKLIFHPQDVPAIRKIREHPAVIDIYEDIDIPAKRY	120
Qy	121	LIDKGLIPWEGDEELKMAFDIETFYHGEDEFKGEIIMISYADESEARVITWKNIDLPY	180	Qy	121	LIDKGLIPWEGDEELKMAFDIETFYHGEDEFKGEIIMISYADESEARVITWKNIDLPY	180
Db	121	LIDKGLVPMWEGDEELKMAFDIETFYHGEDEFKGEIIMISYADESEARVITWKNIDLPY	180	Db	121	LIDKGLVPMWEGDEELKMAFDIETFYHGEDEFKGEIIMISYADESEARVITWKNIDLPY	180
Qy	181	VDVVSHEREMIKFVQVIREKDPDLITVNGNDFDLPYLIKRAEKLGVTLGLGRDKEHPE	240	Qy	181	VDVVSHEREMIKFVQVIREKDPDLITVNGNDFDLPYLIKRAEKLGVTLGLGRDKEHPE	240
Db	181	VDVVSHEREMIKFVQVIREKDPDLITVNGNDFDLPYLIKRAEKLGVTLGLGRDKEHPE	240	Db	181	VDVVSHEREMIKFVQVIREKDPDLITVNGNDFDLPYLIKRAEKLGVTLGLGRDKEHPE	240
Qy	241	PKIHRMGDSFAVEIKGRIHFDLPFVVRRTINPTTYTLEAVYEAVLGKTKLGAEEIAAI	300	Qy	241	PKIHRMGDSFAVEIKGRIHFDLPFVVRRTINPTTYTLEAVYEAVLGKTKLGAEEIAAI	300
Db	239	PKIQMGDRFAVEIKGRIHFDLPFVVRRTINPTTYTLEAVYEAVLGKTKLGAEEITPA	298	Db	239	PKIQMGDRFAVEIKGRIHFDLPFVVRRTINPTTYTLEAVYEAVLGKTKLGAEEITPA	298
Qy	301	WETEESMKLAQYSMEDARATYELGKEFFPMELAKLIGOSVWDVSRSSSTGNLVEWYLL	360	Qy	301	WETEESMKLAQYSMEDARATYELGKEFFPMELAKLIGOSVWDVSRSSSTGNLVEWYLL	360
Db	299	WETGENLVARYSMEDAKVYELGKEFFPMELAKLIGOSVWDVSRSSSTGNLVEWYLL	358	Db	299	WETGENLVARYSMEDAKVYELGKEFFPMELAKLIGOSVWDVSRSSSTGNLVEWYLL	358
Qy	361	RVAYERNELAPNKPDEEYRRLRTTYLGGYVKEPGRGLWENITYLDFRCLYPSIIITHN	420	Qy	361	RVAYERNELAPNKPDEEYRRLRTTYLGGYVKEPGRGLWENITYLDFRCLYPSIIITHN	420
Db	359	RVAYERNELAPNKPDEEYRRLRTTYLGGYVKEPGRGLWENITYLDFRCLYPSIIITHN	417	Db	359	RVAYERNELAPNKPDEEYRRLRTTYLGGYVKEPGRGLWENITYLDFRCLYPSIIITHN	417
Qy	421	VSPDTLREGCKNYDVAPIGVYKFCDFPGFIPSIIGELITMRQBIKKMKATIDPIEKK	480	Qy	421	VSPDTLREGCKNYDVAPIGVYKFCDFPGFIPSIIGELITMRQBIKKMKATIDPIEKK	480
Db	418	VSPDTLNRGCKEYDVAQVGHKFCDFPGFIPSIIGDLLEERQIKKKMKATIDPIEKK	477	Db	418	VSPDTLNRGCKEYDVAQVGHKFCDFPGFIPSIIGDLLEERQIKKKMKATIDPIEKK	477
Qy	481	MLDYRQRAVKLANSYGYMGYPKARWYSKECAESVTAGRWYHIEMTIKEIEEFKGFVL	540	Qy	481	MLDYRQRAVKLANSYGYMGYPKARWYSKECAESVTAGRWYHIEMTIKEIEEFKGFVL	540
Db	478	LLDYRQRAKILANSYGYGYVARAWYCKECAESVTAGREYITMTIKEIEEYGFVVI	537	Db	478	LLDYRQRAKILANSYGYGYVARAWYCKECAESVTAGREYITMTIKEIEEYGFVVI	537
Qy	541	YATDGFYATIPGKPEYTKKAKPELKYINSKLPGLLEVEGYELRGFFVAKRYAVI	600	Qy	541	YATDGFYATIPGKPEYTKKAKPELKYINSKLPGLLEVEGYELRGFFVAKRYAVI	600
Db	538	YSYDGFATIPGADATYKKAMEFLYINAKLPGLLEVEGYELRGFFVAKRYAVI	597	Db	538	YSYDGFATIPGADATYKKAMEFLYINAKLPGLLEVEGYELRGFFVAKRYAVI	597
Qy	601	DEGRITRGLVVRDWSIAKETQAKVLEAILKEDSVKAVEVWQWBEIAKYQVPL	660	Qy	601	DEGRITRGLVVRDWSIAKETQAKVLEAILKEDSVKAVEVWQWBEIAKYQVPL	660
Db	598	DEGKITRGLVVRDWSIAKETQAKVLEAILKEDSVKAVEVWQWBEIAKYQVPL	657	Db	598	DEGKITRGLVVRDWSIAKETQAKVLEAILKEDSVKAVEVWQWBEIAKYQVPL	657
Qy	661	EKLVIHQITKDLSEYKAIQPHVAIAKRLAAGKIKVRPOTIISYIVLKGSGKISRVL	720	Qy	661	EKLVIHQITKDLSEYKAIQPHVAIAKRLAAGKIKVRPOTIISYIVLKGSGKISRVL	720
Db	658	EKLVIHQITRDLKDYKAGPHVAVAKLAAGVIRPGTVISYIVLKGSGRIGRAIPF	717	Db	658	EKLVIHQITRDLKDYKAGPHVAVAKLAAGVIRPGTVISYIVLKGSGRIGRAIPF	717
Qy	721	SEYDPKHKYDPPYIENQVLPVLAIRLAFAGYKEDLAKYQSSKQVGLDAMLK	773	Qy	721	SEYDPKHKYDPPYIENQVLPVLAIRLAFAGYKEDLAKYQSSKQVGLDAMLK	773
Db	718	DEFDPTKHKYDABEYIENQVLPVLAIRLAFAGYKEDLAKYQSSKQVGLDAMLK	770	Db	718	DEFDPTKHKYDABEYIENQVLPVLAIRLAFAGYKEDLAKYQSSKQVGLDAMLK	770
RESULT 2							
US-09-073-354-1							
; Sequence 1, Application US/09073354							
; Patent No. 6033659							
; GENERAL INFORMATION:							
; APPLICANT: KITABAYASHI, Masao							
; APPLICANT: ARAKAWA, Taku							
; APPLICANT: INOUE, Hirotaki							
; APPLICANT: KAWAKAMI, Bunsei							
; APPLICANT: KAWAMURA, Yoshihisa							
; APPLICANT: IMANAKA, Tadayuki							
; APPLICANT: TAKAGI, Masahiro							
; APPLICANT: MORIKAWA, Masaaki							
; TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for							

US-09-073-354-1		Query Match		80.0%; Score 3222.5; DB 3; Length 774;			
		Best Local Similarity		77.4%; Pred. No. 2e-242;			
		Matches 599; Conservative		83; Mismatches 89; Indels 3; Gaps 2;			
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		APPLICANT: INOUE, Hiroaki					
		APPLICANT: KAWAKAMI, Bunsei					
		APPLICANT: KAWAMURA, Yoshihisa					
		APPLICANT: IMANAKA, Tadayuki					
		APPLICANT: TAKAGI, Masahiro					
		APPLICANT: MORIKAWA, Masaaki					
		TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for					
		Sequence 1, Application US/09073354					
		Patent No. 6033859					
		GENERAL INFORMATION:					
		APPLICANT: KITABAYASHI, Masao					
		APPLICANT: ARAKAWA, Taku					
		AP					

Db 418 VSPDTLNREGCKEYDVAPQVGHGFCKPPFPISLLGDLLEERQKIKKKMKATIDPIERK 477
Qy 481 MLDYRQRAVKLHANSYGYGYPKARWYKSCASVTAWGHRHYLWTKIEEEXFGKVL 540
Db 478 LLDYRQRAIKILANSYGYGYPKARWYKSCASVTAWGHRHYLWTKIEEEXFGKVL 537
Qy 541 YADTDGFYATIPGKPEKTIKKAKAEFLKYINSKLPGLLELEYEGFYLRGPFVAKRYAVI 600
Db 538 YSDTDGFTATPGADAETVKKKAMEFLNYINAKLPGALELEYEGFYLRGPFVAKRYAVI 597
Qy 601 DEEGRIITRGLVVRDSEIATKQAVLEAILKEDSVKAEIVKDVVEIAYOVPL 660
Db 598 DEEGRIITRGLVVRDSEIATKQAVLEAILKEDSVKAEIVKDVVEIAYOVPL 657
Qy 661 EKLVIHEQITKOLSEYKAIGPHVAIAKRLAAKGIKVRPOTIISYIVLRGSKISDRVILL 720
Db 658 EKLVIHEQITRDLKDYKATGPHVAVAKRLAAKGIKVRPOTIISYIVLRGSKISDRVILL 717
Qy 721 SEYDPKHKYDPPDYIENQVLPVAILPAFGRKEDLYQSSKQVGLDANLK 773
Db 718 DEFDPKHKYDPPDYIENQVLPVAILPAFGRKEDLYQSSKQVGLDANLK 770

RESULT 3

US-08-656-005A-1
; Sequence 1, Application US/08656005A
; Patent No. 6054301
; GENERAL INFORMATION:
; APPLICANT: KITABAYASHI, Masao
; APPLICANT: ARAKAWA, Taku
; APPLICANT: INOUE, Hiroaki
; APPLICANT: KAWAKAMI, Bunsei
; APPLICANT: KAWAMURA, Yoshihisa
; APPLICANT: IWANAKA, Tadayuki
; APPLICANT: TAKAGI, Masahiro
; APPLICANT: MORIKAWA, Masaaki
; TITLE OF INVENTION: A Method of Amplifying Nucleic
; TITLE OF INVENTION: Acid and A Reagent Therefor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,005A
; FILING DATE: 24 MAY 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 134096/95
; FILING DATE: 31 MAY 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Toffenetti, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 2418/3
; TELEPHONE: 202-429-1776
; TELEFAX: 202-429-0796
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-005A-1

Query Match 80.0%; Score 3222.5; DB 3; Length 774;
Best Local Similarity 77.4%; Pred No. 2a-242;
Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

Qy 1 MIPDTYITKOKPIIRIPFKKENGFBKIELDPHFOPYIYALLKDDSAIDIEIKAIKERHG 60
Db 1 MILDYITIEDGKPVIRIPFKKENGFBKIEYDRTFSPFYALLKDDSAIDIEIKAIKERHG 60
Qy 61 KIVRVVDAVKKFLGRDVEVWKLIFHPQDPVAPKIRREHPAVIDIYVDIPAKRY 120
Db 61 TVTVRVKRVKVKFLGRDVEVWKLIFHPQDPVAPKIRREHPAVIDIYVDIPAKRY 120
Qy 121 LIDKGLIPMEGDEELKMAFDIETFYHGEDEFGKEIIMISYADSEEAARVITWKNIDLPY 180
Db 121 LIDKGLVPMGEDEELKMAFDIETFYHGEDEFGKEIIMISYADSEEAARVITWKNIDLPY 180
Qy 181 DVVSVNERMIKRFVQIVREKDPDLITVNGDNFDPYILKRAEKLGVTLILGRDKEHE 240
Db 181 DVVSVNERMIKRFVQIVREKDPDLITVNGDNFDPYILKRAEKLGVTLILGRDKEHE 240
Qy 241 PKIHRMGDSFAVEIKGRIHFDLPFVVRRTINLPTTYLEAVYEAIVLGKTKSLGABEIAA 300
Db 239 PKIHRMGDSFAVEIKGRIHFDLPFVVRRTINLPTTYLEAVYEAIVLGKTKSLGABEIAA 298
Qy 301 WETERSMKKLAQYSMEDAZATYELGKEFPPEMAELAKLIGOSVWDSRSTGNLWVILL 360
Db 299 WETGENLVARYSMEDAKVTYELGKEFLPMBAQUSRLIGOSLMDVSRSTGNLWVILL 358
Qy 361 RVAYERNELAPNKPDEEYRRRLRTTYLGGYVKEPERGLWENITTYLDFCLYPSIIVTHN 420
Db 359 RVAYERNELAPNKPDEEYRRRLRTTYLGGYVKEPERGLWENITTYLDFCLYPSIIVTHN 417
Qy 421 VSPDTLEREGCKYDVAPIVGVKFCDFPFPISLLGDLLEERQKIKKKMKATIDPIERK 480
Db 418 VSPDTLEREGCKYDVAPIVGVKFCDFPFPISLLGDLLEERQKIKKKMKATIDPIERK 477
Qy 481 MLDYRQRAVKLHANSYGYGYPKARWYKSCASVTAWGHRHYLWTKIEEEXFGKVL 540
Db 478 LLDYRQRAIKILANSYGYGYPKARWYKSCASVTAWGHRHYLWTKIEEEXFGKVL 537
Qy 541 YADTDGFYATIPGKPEKTIKKAKAEFLKYINSKLPGLLELEYEGFYLRGPFVAKRYAVI 600
Db 538 YSDTDGFTATPGADAETVKKKAMEFLNYINAKLPGALELEYEGFYLRGPFVAKRYAVI 597
Qy 601 DEEGRIITRGLVVRDSEIATKQAVLEAILKEDSVKAEIVKDVVEIAYOVPL 660
Db 598 DEEGRIITRGLVVRDSEIATKQAVLEAILKEDSVKAEIVKDVVEIAYOVPL 657
Qy 661 EKLVIHEQITKOLSEYKAIGPHVAIAKRLAAKGIKVRPOTIISYIVLRGSKISDRVILL 720
Db 658 EKLVIHEQITRDLKDYKATGPHVAVAKRLAAKGIKVRPOTIISYIVLRGSKISDRVILL 717
Qy 721 SEYDPKHKYDPPDYIENQVLPVAILPAFGRKEDLYQSSKQVGLDANLK 773
Db 718 DEFDPKHKYDPPDYIENQVLPVAILPAFGRKEDLYQSSKQVGLDANLK 770

RESULT 4

US-09-073-259-1
; Sequence 1, Application US/09073259
; Patent No. 6143536
; GENERAL INFORMATION:
; APPLICANT: IMANAKA, Tadayuki
; APPLICANT: TAKAGI, Masahiro
; APPLICANT: MORIKAWA, Masaaki
; TITLE OF INVENTION: DNA Encoding A Thermostable DNA Polymerase
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.

```

/ COUNTRY: USA
/ ZIP: 20036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WordPerfect 6.1 Windows
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/073,259
/ FILING DATE: Concurrent Herewith
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/656,005
/ FILING DATE: 24 MAY 1996
/ APPLICATION NUMBER: JP 134096/95
/ FILING DATE: 31 MAY 1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Toffenetti, Judith L.
/ REGISTRATION NUMBER: 39,048
/ REFERENCE/DOCKET NUMBER: 2418/10
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-429-1776
/ TELEFAX: 202-429-0796
/ INFORMATION FOR SEQ ID NO. 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 774 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-073-259-1

Query Match 80.0%; Score 3222.5; DB 3; Length 774;
Best Local Similarity 77.4%; Pred. No. 2e-242;
Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

QY 1 MIFDTYITKDGKPIIRIFKKGEGFKIELDPHFQPIYIALLKDDSAIDRIKAIKGERHG 60
DB 1 MILDYITEDGKPIIRIFKKGEGFKIELDRTPEFYALLKDDSAIEVVKITAERHG 60
QY 61 KIVRVVDAVKVKKFLGRDVEVWKLIFEHPQDVPAALRGKIREHPAVIDIYEDIPFAKRY 120
DB 61 TVTVTKRVEKVKKFLGRDVEVWKLIFEHPQDVPAALRGKIREHPAVIDIYEDIPFAKRY 120
QY 121 LIDKGLIPMEGDEELKMAFDIETPYHEGDFGKEIIMISYADEEGARVITWKNIDLPY 180
DB 121 LIDKGLVPMGDEELKMAFDIOTLYHEGEFAEGPILMISYADEEGARVITWKNVDLPY 180
QY 181 VDVSNEREMIKRFVQIVREKDPDVLITYNGDNFDLPYLKKEAKLGVTLILGRDKEHPE 240
DB 181 VDVSNEREMIKRFVQIVREKDPDVLITYNGDNFDLPYLKKEAKLGVTLILGRDKEHPE 240
QY 241 PKIHRMGDSFAVEIKGRIFHFDLPVVRRTINLPTTYLLEAVYEAVALGKTSKLGAEETAAI 300
DB 239 PKIQRMGDRFAVEIKGRIFHFDLPVVRRTINLPTTYLLEAVYEAVALGKTSKLGAEETAAI 300
QY 301 WETESMKLAQYSMEDARATYELGKEFFPMEAEALKIGOSVNDVSRSSSTGNLVEWVLL 360
DB 299 WETGENLERVARYSMEDAKVTYELGKEFFPMEAEALKIGOSVNDVSRSSSTGNLVEWVLL 360
QY 361 RVAYERNELAPNKPDDEEYRRRLRTTYLGGYVKEPERGLWENITYLDFRCLYPSIIIVTHN 420
DB 359 RKAYERNELAPNKPDDEEYRRRLRTTYLGGYVKEPERGLWENITYLDFRCLYPSIIIVTHN 420
QY 421 VSPDTLNRGCKEYDVAPOVGHFCKDPFGFIPSLGDLLEERQIKKKMATTIDPIERK 480
DB 418 VSPDTLNRGCKEYDVAPOVGHFCKDPFGFIPSLGDLLEERQIKKKMATTIDPIERK 480
QY 481 MLDYRQRAVKLHANSYGYGVKPKARWYSKECAESVTAWGRHYITMTIKETEEKFGKVL 540
DB 478 LLDYRQRAIKILANSYGYGVKPKARWYSKECAESVTAWGRHYITMTIKETEEKFGKVL 540
QY 541 YADTDGFFATIPGADAEVTKKAMEFLNYNAKLPALBLEYEGFYKRGFFVTKKKYAVI 597
DB 538 YSDTDGFFATIPGADAEVTKKAMEFLNYNAKLPALBLEYEGFYKRGFFVTKKKYAVI 597
QY 601 DEEGRIITRGLVVRDWSIAKETQAKVLEAILKEDSVKEAVEIKVDVVBEIAKQVPL 660
DB 598 DEEGRIITRGLVVRDWSIAKETQAKVLEAILKEDSVKEAVEIKVDVVBEIAKQVPL 660
QY 661 EKLVIHQITKDLSEYKAIQPHVAIAKRLAAKIKVRPGTIISYIVLRSGKISDRVILL 720
DB 658 EKLVIHQITKDLSEYKAIQPHVAIAKRLAAKIKVRPGTIISYIVLRSGKISDRVILL 720
QY 721 SEYDPKXHKYDPPYIENQVLPVLRILZAFGYKEDKLYQSSKQVGLDAMLK 773
DB 718 DEFDPKXHKYDPPYIENQVLPVLRILZAFGYKEDKLYQSSKQVGLDAMLK 773

RESULT 5
US-09-363-095-1
; Sequence 1, Application US/09363095
; Patent No. 6187573
; GENERAL INFORMATION:
; APPLICANT: IMANAKA, Tadayuki
; APPLICANT: TAKAGI, Masahiro
; APPLICANT: MORIKAWA, Masaaki
; TITLE OF INVENTION: DNA Encoding A Thermostable DNA Polymerase
; FILE REFERENCE: 2418/11
; CURRENT APPLICATION NUMBER: US/09/363,095
; CURRENT FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: US 09/073,259
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: US 08/656,005
; EARLIER FILING DATE: 1996-05-24
; EARLIER APPLICATION NUMBER: JP 134096/95
; EARLIER FILING DATE: 1995-05-31
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: WordPerfect 6.1 Windows
; SEQ ID NO 1
; TYPE: PRT
; LENGTH: 774
; ORGANISM: Hyperthermophilic archaeon
; US-09-363-095-1

Query Match 80.0%; Score 3222.5; DB 3; Length 774;
Best Local Similarity 77.4%; Pred. No. 2e-242;
Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

QY 1 MIFDTYITKDGKPIIRIFKKGEGFKIELDPHFQPIYIALLKDDSAIDRIKAIKGERHG 60
DB 1 MILDYITEDGKPIIRIFKKGEGFKIELDRTPEFYALLKDDSAIEVVKITAERHG 60
QY 61 KIVRVVDAVKVKKFLGRDVEVWKLIFEHPQDVPAALRGKIREHPAVIDIYEDIPFAKRY 120
DB 61 TVTVTKRVEKVKKFLGRDVEVWKLIFEHPQDVPAALRGKIREHPAVIDIYEDIPFAKRY 120
QY 121 LIDKGLIPMEGDEELKMAFDIETPYHEGDFGKEIIMISYADEEGARVITWKNIDLPY 180
DB 121 LIDKGLVPMGDEELKMAFDIOTLYHEGEFAEGPILMISYADEEGARVITWKNVDLPY 180
QY 181 VDVSNEREMIKRFVQIVREKDPDVLITYNGDNFDLPYLKKEAKLGVTLILGRDKEHPE 240
DB 181 VDVSNEREMIKRFVQIVREKDPDVLITYNGDNFDLPYLKKEAKLGVTLILGRDKEHPE 240
QY 241 PKIHRMGDSFAVEIKGRIFHFDLPVVRRTINLPTTYLLEAVYEAVALGKTSKLGAEETAAI 300
DB 239 PKIQRMGDRFAVEIKGRIFHFDLPVVRRTINLPTTYLLEAVYEAVALGKTSKLGAEETAAI 300
QY 301 WETESMKLAQYSMEDARATYELGKEFFPMEAEALKIGOSVNDVSRSSSTGNLVEWVLL 360
DB 299 WETGENLERVARYSMEDAKVTYELGKEFFPMEAEALKIGOSVNDVSRSSSTGNLVEWVLL 360
QY 361 RVAYERNELAPNKPDDEEYRRRLRTTYLGGYVKEPERGLWENITYLDFRCLYPSIIIVTHN 420
DB 359 RKAYERNELAPNKPDDEEYRRRLRTTYLGGYVKEPERGLWENITYLDFRCLYPSIIIVTHN 420
QY 421 VSPDTLNRGCKEYDVAPOVGHFCKDPFGFIPSLGDLLEERQIKKKMATTIDPIERK 480
DB 418 VSPDTLNRGCKEYDVAPOVGHFCKDPFGFIPSLGDLLEERQIKKKMATTIDPIERK 480
QY 481 MLDYRQRAVKLHANSYGYGVKPKARWYSKECAESVTAWGRHYITMTIKETEEKFGKVL 540
DB 478 LLDYRQRAIKILANSYGYGVKPKARWYSKECAESVTAWGRHYITMTIKETEEKFGKVL 540
QY 541 YADTDGFFATIPGADAEVTKKAMEFLNYNAKLPALBLEYEGFYKRGFFVTKKKYAVI 600
DB 538 YSDTDGFFATIPGADAEVTKKAMEFLNYNAKLPALBLEYEGFYKRGFFVTKKKYAVI 600

```

```

Qy 421 VSPDTEREGCKNYDVAPIGVYKFCDFPGFIPSLIGELITMQEIKKKOMKATIDPIEK 480
Db 418 VSPDTEREGCKNYDVAPIGVYKFCDFPGFIPSLIGELITMQEIKKKOMKATIDPIEK 477
Qy 481 MLDYRQRAVKLHANSYGYGMYGPKARWYKCECAESVTAWGRHYIEMTIKEIEKFGPKVL 540
Db 478 LLDYRQRAIKLANSYGYGYARARWYKCECAESVTAWGRHYIEMTIKEIEKFGPKVI 537
Qy 541 YADTDGYATIPGKPTTIKKAKFELKYNKSLPGLLELEYEGFYLRGFFVAKKRYAVI 600
Db 538 YSDTDGFFATIPGADAETVKKAMEFLYNINAKLPGLALELEYEGFYLRGFFVAKKRYAVI 597
Qy 601 DEGRITTRGLEIVRRDWSIAKETQAKVLEAILKEDSVKAEIVKDVVEETAKYQVPL 660
Db 598 DEGRITTRGLEIVRRDWSIAKETQAKVLEAILKEDSVKAEIVKDVVEETAKYQVPL 657
Qy 661 EKLVIHQITKDLSEYKAIGHVAIAKRLAAKGIKVRPGTIISYIVLRSGKISDRVILL 720
Db 658 EKLVIHQITRDLKDYKATGPHVAVAKRLAARGVKIRPGTVISYIVLRSGKISDRVILL 717
Qy 721 SEYDPKXHKYDPPDYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAWLK 773
Db 718 DEFDPKHKYDAEYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAWLK 770

```

RESULT 6

```

US-09-418-027-1
; Sequence 1, Application US/09418027
; Patent No. 6225065
; GENERAL INFORMATION:

```

```

; APPLICANT: KITABAYASHI, Masao
; APPLICANT: ARAKAWA, Taku
; APPLICANT: INOUE, Hiroaki
; APPLICANT: KAWAKAMI, Bunsei
; APPLICANT: KAWAMURA, Yoshihisa
; APPLICANT: IMANAKA, Tadayuki
; APPLICANT: TAKAGI, Masahiro
; APPLICANT: MORIKAWA, Masaaki

```

```

; TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for
; AMPLIFYING NUCLEIC ACIDS

```

```

; NUMBER OF SEQUENCES: 16

```

```

; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Kenyon & Kenyon

```

```

; STREET: 1025 Connecticut Avenue, N.W., Suite 600

```

```

; CITY: Washington

```

```

; STATE: D.C.

```

```

; COUNTRY: U.S.A.

```

```

; ZIP: 20036

```

```

; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: 3 5" floppy disk

```

```

; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: WordPerfect 6.1 Windows

```

```

; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/418,027

```

```

; FILING DATE:

```

```

; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 09/073,354

```

```

; FILING DATE:

```

```

; APPLICATION NUMBER: JP 134096/95

```

```

; FILING DATE: 31 MAY 1995

```

```

; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Toffenetti, Judith L.

```

```

; REGISTRATION NUMBER: 39,048

```

```

; REFERENCE/DOCKET NUMBER: 2418/9

```

```

; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 202-429-1776

```

```

; TELEFAX: 202-429-0796

```

```

; INFORMATION FOR SEQ ID NO: 1:

```

```

; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 774 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-418-027-1

```

```

Query Match 80.0%; Score 3222.5; DB 3; Length 774;
Best Local Similarity 77.4%; Pred. No. 2e-242;
Matches 598; Conservative 83; Mismatches 99; Indels 3; Gaps 2;

```

```

Qy 1 MIFDIDYTKDGKPIIRIFKKEGFEKIELDPHFQPIYVALLKDDSAIDEIKAIKGERHG 60
Db 1 MLDTDYITEDGKPVIRIFKKEGFEKIELDPHFQPIYVALLKDDSAIDEIKAIKGERHG 60
Qy 61 KIVRVDAVKKPLGRDVEVWKLIFEHPODVPAALRGKIREHPAVIDIYDIYDIPFAKRY 120
Db 61 TVTVTKRVKVKQKPLGRDVEVWKLIFEHPODVPAALRGKIREHPAVIDIYDIYDIPFAKRY 120
Qy 121 LIDKGLIPMEGDEELKMAFDIETFYHEGDFGKEIIMISYADEEARVITWKNIDLPY 180
Db 121 LIDKGLIPMEGDEELKMAFDIETFYHEGDFGKEIIMISYADEEARVITWKNIDLPY 180
Qy 181 VDVVSNREMIKRFVQIVREKDPDLITVYNGDNFDLPYLKRAEKLGVTLGLGRDKEHPE 240
Db 181 VDVVSTEREMIKRFLRVVVKEXDPDLITVYNGDNFDLPYLKRAEKLGVTLGLGRDKEHPE 240
Qy 241 PKIHRMGDSFAVEIKGRITHFDLPVVRTINLPVTLVAVYVAVLTKSKLGAEEIAAI 300
Db 239 PKIHRMGDSFAVEIKGRITHFDLPVVRTINLPVTLVAVYVAVLTKSKLGAEEIAAI 298
Qy 301 WETESMKKLAQYSMEDARATYELGKFFPMEAEALAKLIGQSVWDVSRSTGNLVEWYLL 360
Db 299 WETGENLVARVYSMEDAKVYELGKFFPMEAEALAKLIGQSVWDVSRSTGNLVEWYLL 358
Qy 361 RVAVERNELAPNKDEBEYERELRTVYLGKVPKPERGLMENITVLPCLYPSIIVTHN 420
Db 359 RVAVERNELAPNKDEBEYERELRTVYLGKVPKPERGLMENITVLPCLYPSIIVTHN 417
Qy 421 VSPDTEREGCKNYDVAPIGVYKFCDFPGFIPSLIGELITMQEIKKKOMKATIDPIEK 480
Db 418 VSPDTEREGCKNYDVAPIGVYKFCDFPGFIPSLIGELITMQEIKKKOMKATIDPIEK 477
Qy 481 MLDYRQRAVKLHANSYGYGMYGPKARWYKCECAESVTAWGRHYIEMTIKEIEKFGPKVL 540
Db 478 LLDYRQRAIKLANSYGYGYARARWYKCECAESVTAWGRHYIEMTIKEIEKFGPKVI 537
Qy 541 YADTDGYATIPGKPTTIKKAKFELKYNKSLPGLLELEYEGFYLRGFFVAKKRYAVI 600
Db 538 YSDTDGFFATIPGADAETVKKAMEFLYNINAKLPGLALELEYEGFYLRGFFVAKKRYAVI 597
Qy 601 DEGRITTRGLEIVRRDWSIAKETQAKVLEAILKEDSVKAEIVKDVVEETAKYQVPL 660
Db 598 DEGRITTRGLEIVRRDWSIAKETQAKVLEAILKEDSVKAEIVKDVVEETAKYQVPL 657
Qy 661 EKLVIHQITKDLSEYKAIGHVAIAKRLAAKGIKVRPGTIISYIVLRSGKISDRVILL 720
Db 658 EKLVIHQITRDLKDYKATGPHVAVAKRLAARGVKIRPGTVISYIVLRSGKISDRVILL 717
Qy 721 SEYDPKXHKYDPPDYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAWLK 773
Db 718 DEFDPKHKYDAEYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAWLK 770

```

RESULT 7

```

US-08-906-925-4
; Sequence 4, Application US/08906925
; Patent No. 5882904
; GENERAL INFORMATION:

```

```

; APPLICANT: Riedl, William

```

```

; APPLICANT: Fly, Susan

```

```

; APPLICANT: Kaboord, Barbara F.

```

```

; APPLICANT: Nye, Steven H.

```

```

; APPLICANT: Ting, Eve

```

;; TITLE OF INVENTION: THERMOCOCCUS BAROSII DNA POLYMERASE MUTANTS
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Quarles & Brady
;; STREET: 411 East Wisconsin Avenue
;; CITY: Milwaukee
;; STATE: Wisconsin
;; COUNTRY: U.S.A.
;; ZIP: 53202-4497
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/906,925
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Baker, Jean C.
;; REGISTRATION NUMBER: 35,433
;; REFERENCE/DOCKET NUMBER: 740211.90628
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (414) 277-5709
;; TELEFAX: (414) 271-3552
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 778 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-906-925-4

Query Match 79.7%; Score 3209.5; DB 2; Length 778;
Best Local Similarity 77.2%; Pred. No. 2.1e-241;
Matches 596; Conservative 93; Mismatches 90; Indels 3; Gaps 2;
QY 1 MIFDYYITKDGKPIIRIFKXGKGFKIELDPHFQPIYVALLKDSALDEIKAKGERHG 60
DB 1 MILVDYITEDGKPIRVFKDKGFKLEYDREFEPYIYALLRDSALDEIEKITAERHG 60
QY 61 KIVRVDAVKVKKGLGRVVEWKLIFHPQVDPALRGKIRHPAVIDIYDIPFAKRY 120
DB 61 KVVVKRAEKVKKGLGRSEVWVLYFHPQVPAIRDKIRHPAVIDIYDIPFAKRY 120
QY 121 LIDKGLIPMEGDEELKMAFDIETPVHGEDEFGKEIIMISYADDEEARVITWKNIDLPY 180
DB 121 LIDKGLVPNEGDEELKMSFDIETLYHEGEERGTGPILMISYADSEARVITWKNIDLPY 180
QY 181 VDVSNEREMIKRFVQIVREKDPDLITVNGDNFDPYLIIRAEKGLVTLILGRDKEHPE 240
DB 181 VDWVSTEKEMIKRFLKVVKEKDPDLITVNGDNFDPYLIIRAEKGLVSTLGRDGS--E 238
QY 241 PKIHRMGDSFAVEIKRHFDPVVRRTINLPTVTLAVYBAVLGKTKSLGABEIAAI 300
DB 239 PKIQMGDRFAVEVKRHFDPVVRRTINLPTVTLAVYBAVFGKPKVKYAEIATA 298
QY 301 WTEESMKLAQYSMEDARATYELGKEFPMPMAELAKIGQSVWDVSRSSSTGNLVWYLL 360
DB 299 WETGELERVARYSMEDARATYELGKEFPMPMAEQSLRIGQGLWDVSRSSSTGNLVWELL 358
QY 361 RVAYERNELAPNKPDEEYRRRLRTYLGYYKKEPERGLWENITVLDRCYPSIIVTHN 420
DB 359 RYAYERNELAPNKPDEELARR--RGYAGYKKEPERGLWMDVILDFRSYPSIIVTHN 417
QY 421 VSPDTLEREGCKNYDVAPIVGYKFKDPCFPFIPSLGELITMRQBIKKMKATIDPIBK 480
DB 418 VSPDTLNRGCKSYDVAPQVGHKFKDPCFPFIPSLGELNLEERQIKRKMATLDPLERK 477
QY 481 MLDYRORAVKLHANSYGYWGPYKARYSKCEASVTWANGRYIEWTIKEEKEGFKVL 540
DB 478 LLDYRORAILANSFYGYGYARWYKCEASVTWANGRYIEWIREEKEGFKVL 537

QY 541 YATDGFYATIPGKPEKITKKAKPELKYINKLPGLLELEYEGFYLRGFFVAKRYAVI 600
DB 538 YATDGLHATIFGADAEIVKKKXAMFELNINPKLPGLLELEYEGFYVRGFFVTKKYAVI 597
QY 601 DEGRITTRGLEVVRDSEIIEAKETQAKVLEBAILKEDSVKAVEIVKDVVEIAKYQVFL 660
DB 598 DEEGKITTRGLEIVREDSEIIEAKETQARVLEAILRHGDVEEAVRIYKVEKLSKYEVPP 657
QY 661 EKLVIHEQITKDLSEYKAIGPHVAIAKRLAAKIGKVRPGTIISYIVLRGSGKISDRVILL 720
DB 658 EKLVIHEQITRELKDYKATGPHVAIAKRLAARGIKIRPGTVISYIVLRGSGRIGRAIFP 717
QY 721 SEYDPKHKYDPPYIENQVLPVAVURILEAFGRKEDLKYQSSKQVGLDAWL 772
DB 718 DEEDPTKHYDADYIENQVLPVAVURILEAFGRKEDLKYQTRQVGLGAWL 769
RESULT 8
US-07-966-278-1
; Sequence 1, Application US/07966278
; Patent No. 5489523
; GENERAL INFORMATION:
; APPLICANT: Mathur, Eric A
; TITLE OF INVENTION: EXONUCLEASE-DEFICIENT THERMOSTABLE
; TITLE OF INVENTION: PYROCOCUS FURIOSUS DNA POLYMERASE I
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Road, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/966,278
; FILING DATE: 19921226
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: STG0133P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-07-966-278-1

Query Match 78.3%; Score 3151; DB 1; Length 775;
Best Local Similarity 74.8%; Pred. No. 7.4e-237;
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;
QY 1 MIFDYYITKDGKPIIRIFKXGKGFKIELDPHFQPIYVALLKDSALDEIKAKGERHG 60
DB 1 MILVDYITTEGKPIRVLFKXGKGFKIEHDTFRFYIYALLRDSKIEEVKKITGERHG 60
QY 61 KIVRVDAVKVKKGLGRVVEWKLIFHPQVDPALRGKIRHPAVIDIYDIPFAKRY 120

Db 61 KIVRVDVEKVEKFLGKPIVWKLYLHPQDVPTIREKVRHPAVVDIFEYDIPFAKRY 120
Qy 121 LIDKGLIPMEGBELKLMADFIETFYHEGDEFGKGEIIMISYADEBEARVITWKNIDLPY 180
Db 121 LIDKGLIPMEGBELKLMADFIETFYHEGDEFGKGEIIMISYADEBEARVITWKNIDLPY 180
Qy 181 VDVVSNREMIKFPVQIVREKDPDLITYNGDNFDPYLIKRAEKIGVTLGLGRDKEHPE 240
Db 181 VEVVSSREMIKFLRIIREKDPDIIVYNGDSFDPYLAKEKRAEKLIGIKLTIGRDS--E 238
Qy 241 PKIHRMGDSFAVEIKGRIHFDLPVVRRTINLPTYTLEAVYEAVLGKTSKLGAEIAAI 300
Db 239 PKMQRIGDMTAVEVKGRIHFDLYHVITRITNLTPTYLEAVYEAIFGKPKKVVADSIKA 298
Qy 301 WETEESMKLAQYSMEDARATYELGKEFPFMEAEALAKLGOSVWDVSRSTGNLVEWYLL 360
Db 299 WESGENLVRVAKYSMEDAKATYELGKEFLPMETQLSRLVGQPLWDVSRSTGNLVEWFL 358
Qy 361 RVAYERNELAPNKPDEEYRRRTTIVGKYVKEPERGLWENITYLDFCLYPSIIVTN 420
Db 359 RKAYERNEVAPNKPSEEEYQRRRESYTGFGVKEPEKGLWENIVYLDFFALYPSIITHN 418
Qy 421 VSPDTLEREGCKNYDVAPIVGYKCDPFGFIPSIILGELITMRQEIKKWKATIDPIEK 480
Db 419 VSPDTLNLGCKNYDIAPOVGHKFCDDIPGFIPLSLGLHLEERQIKTKWKETQDPIEK 478
Qy 481 MLDYRQRAVKLHANSYGYMGYPKARWYSKECAESVTANGRHYEMTIEIEBKFGFKVL 540
Db 479 LLDYRQKAIKLANSPYGYGAKARWYCKECAESVTANGRKYIELWKELEBKFGFKVL 538
Qy 541 YADTDGYATIPGKEPETTKKAKPELKYNSKLPGLLELYEGFYLRGFFVAKRYAVI 600
Db 539 YIDTDGLYATIPGSESEIKKALEFPVKYNSKLPGLLELYEGFYLRGFFVTKRYAVI 598
Qy 601 DEGRITTRGLEVRDWSIEIAKETQAKVLEBALKEDSVKEAVEIVKDWEEIAKYQVPL 660
Db 599 DEBGKITRGLEVRDWSIEIAKETQARVLETILRHGDVEEAIRIVKEVIQKLANYEIPP 658
Qy 661 EKLVIHQITKOLSEYKAIQPHVAIAKRAAAGIKVRPGTIISYIVLRSGKISDRVILL 720
Db 659 EKLAIEQITRPLEYKAIQPHVAVAKLAAGKVKIPGWYIGYVLRGDSINRAIIA 718
Qy 721 SEYDPKHKYDDPYIENQVLEPAVLEAFGRYKEDLKYQSKQVGLDAWL 772
Db 719 BEYDPKHKYDABEYIENQVLEPAVLEAFGRYKEDLRYQKTRQVGLTSL 770

RESULT 9

US-08-424-921-1
; Sequence 1: Application US/08424921
; Patent No. 5545552
; GENERAL INFORMATION:
; APPLICANT: Mathur, Eric A
; TITLE OF INVENTION: PURIFIED THERMOSTABLE PYROCOCUS
; TITLE OF INVENTION: FURIOSUS DNA POLYMERASE I
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bingham & Pitting
; STREET: 12526 High Bluff Road, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC Compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,921
; FILING DATE: 19-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/803,627
; FILING DATE: 02-DEC-1991
; APPLICATION NUMBER: US 07/620,568
; FILING DATE: 03-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,073
; FILING DATE: 19-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/776,552
; FILING DATE: 15-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: STG01COP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-424-921-1

Query Match 78.3%; Score 3151; DB 1: Length 775;
Best Local Similarity 74.6%; Pred. No. 7.4e-937;
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

Qy 1 MIFDTYITDKGPIITRIFKENGFEKIELDPHFQPIYIALLKDDSAIDEIKAKGRHG 60
Db 1 MILDVDYITBEGFVIRLFKENGKFKIEHRTFRPIYIALLRDDSKIEBVKKITGRHG 60
Qy 61 KIVRVDPAVKVKKFLGRDVEVWKLIFEPHQDVPALRGKIREHPAVDIFEYDIPFAKRY 120
Db 61 KIVRVDVEKVEKFLGKPIVWKLYLHPQDVPTIREKVRHPAVVDIFEYDIPFAKRY 120
Qy 121 LIDKGLIPMEGBELKLMADFIETFYHEGDEFGKGEIIMISYADEBEARVITWKNIDLPY 180
Db 121 LIDKGLIPMEGBELKLMADFIETFYHEGDEFGKGEIIMISYADEBEARVITWKNIDLPY 180
Qy 181 VDVVSNREMIKFPVQIVREKDPDLITYNGDNFDPYLIKRAEKIGVTLGLGRDKEHPE 240
Db 181 VEVVSSREMIKFLRIIREKDPDIIVYNGDSFDPYLAKEKRAEKLIGIKLTIGRDS--E 238
Qy 241 PKIHRMGDSFAVEIKGRIHFDLPVVRRTINLPTYTLEAVYEAVLGKTSKLGAEIAAI 300
Db 239 PKMQRIGDMTAVEVKGRIHFDLYHVITRITNLTPTYLEAVYEAIFGKPKKVVADSIKA 298
Qy 301 WETEESMKLAQYSMEDARATYELGKEFPFMEAEALAKLGOSVWDVSRSTGNLVEWYLL 360
Db 299 WESGENLVRVAKYSMEDAKATYELGKEFLPMETQLSRLVGQPLWDVSRSTGNLVEWFL 358
Qy 361 RVAYERNELAPNKPDEEYRRRTTIVGKYVKEPERGLWENITYLDFCLYPSIIVTN 420
Db 359 RKAYERNEVAPNKPSEEEYQRRRESYTGFGVKEPEKGLWENIVYLDFFALYPSIITHN 418
Qy 421 VSPDTLEREGCKNYDVAPIVGYKCDPFGFIPSIILGELITMRQEIKKWKATIDPIEK 480
Db 419 VSPDTLNLGCKNYDIAPOVGHKFCDDIPGFIPLSLGLHLEERQIKTKWKETQDPIEK 478
Qy 481 MLDYRQRAVKLHANSYGYMGYPKARWYSKECAESVTANGRHYEMTIEIEBKFGFKVL 540
Db 479 LLDYRQKAIKLANSPYGYGAKARWYCKECAESVTANGRKYIELWKELEBKFGFKVL 538
Qy 541 YADTDGYATIPGKEPETTKKAKPELKYNSKLPGLLELYEGFYLRGFFVAKRYAVI 600
Db 539 YIDTDGLYATIPGSESEIKKALEFPVKYNSKLPGLLELYEGFYLRGFFVTKRYAVI 598
Qy 601 DEGRITTRGLEVRDWSIEIAKETQAKVLEBALKEDSVKEAVEIVKDWEEIAKYQVPL 660
Db 719 BEYDPKHKYDABEYIENQVLEPAVLEAFGRYKEDLRYQKTRQVGLTSL 770

Db 599 DEEGKVTITRGLIIVRRDWSIAKETQARVLETLKHGDEVEAVRIIVKEVIQKLANYEIPP 658
Qy 661 EKLVIHQITKDSYKAIQPHVAIAKRLAAKGIKVRPGTIISYIVLRSGKTSRDRVILL 720
Db 659 EKLAIVEQITRPLHEYKAIGPHVAVAKKLAAGVKIKPGMVGIVYIVLRGDPISNRILA 718
Qy 721 SEYDPKHKYDPPDYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAML 772
Db 719 EYDPPKHKYDAEYIENQVLPVLRILEFGYRKEDLRYQKTRQVGLTSM 770

RESULT 10
US-08-556-355A-1
; Sequence 1, Application US/08556355A
; Patent No. 5966395
; GENERAL INFORMATION:
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: Purified Thermostable Pyrococcus
; TITLE OF INVENTION: furiosus DNA Polymerase I
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556.355A
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/424,921
; FILING DATE: 19-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,627
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,846
; FILING DATE: 21-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,073
; FILING DATE: 19-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,568
; FILING DATE: 03-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. Paul
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 04121.0004-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-556-355A-1

Query Match 78.3%; Score 3151; DB 2; Length 775;
Best Local Similarity 74.6%; Pred. NO. 7.4e-237;
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

Qy 1 MIFDVTITKDGKPIIRIFKENGEPKIELDPHFQPIYALKKDSIDAIDEIKAIKGERHG 60

Db 1 MILDVDYITTEGKPVIRLFKEKNGFKLEHDTFRPYIALLRDDSKLEEYKKITGERHG 60
Qy 61 KIVRVVDVAVKVKFLGRDVEVWKLIFEPHODVPALRGKIREHPAVDIIEYDIPFAKRY 120
Db 61 KIVRIVDVEKVEKFLGKPIITWKLYLEHPQDVPTIREKVRHPAVVDIIEYDIPFAKRY 120
Qy 121 LIDKGLIPMEGDEELKLMAFDIETFYHEGDFGKEIIMISYADEEEARVITWKNIDLPY 180
Db 121 LIDKGLIPMEGDEELKLMAFDIETFYHEGDFGKEIIMISYADEEEARVITWKNIDLPY 180
Qy 181 VDVSNREMIKRFVQIVREKDPDVLITYNQDNFDLPYLIKRAEKLGVTLILLGRDKEHPE 240
Db 181 VEVSSSEREMIKRFLIRIREKDPDIITVYNGDSDFPYLAKRAEKLGITKITIGSDGS--E 238
Qy 241 PKIHEMGDSFAVEIKGRITHPDLFPVVRTINLPYTLEAVVEAVLGTKSKLGAEEAAAI 300
Db 239 PMQRIQDMTAVEVKGRIHFDLYHVIPTINLPYTLEAVVEAVLGTKSKLGAEEAAAI 298
Qy 301 METESMKLAQYSMEDARATYELGKEFPMEABELAKIIGQSVMDVRSSTGNLVEWYLL 360
Db 299 MESGENLERVAKYSMEDAKATYELGKEFPMEIQLSRLVGQFLMDVRSSTGNLVEWYLL 358
Qy 361 RVAYERNELAPNKDEEYRRLRTTYLGGVVKPERGLWENITVDFRCLYPSIIIVTHN 420
Db 359 RKAYERNEVAPNKPSESEYQRLRESYTGCGFKPEKGLWENIYVDFRALYPSIIIVTHN 418
Qy 421 VSPDTLREGCKNYDVAPIVGYKFCOKPPGPIPSILGELITWRQEIKKOKKATIDPIBK 480
Db 419 VSPDTLNLGCKNYDIAPOVGHKFKCDIPGPIPSILGHLLEERQKIKTKMKETODPIKI 478
Qy 481 MLDYRQRAVKLHANSYGYMGYPKARYSKCARSVTAGRHYIEMTKETEKEFGKVL 540
Db 479 LLDYRQKAIKLANSFYGYGAKARWYCKEARSVTAGRKYIELVWKELEKEFGKVL 538
Qy 541 YADTDGFVATIPGKPEPIKKAKEFLKYINSKLPGLLELEYEGFYLGGFFVAKRYAVI 600
Db 539 YIDTDGLYATIPGSESEIKKALEFVKYINSKLPGLLELEYEGFYLGGFFVAKRYAVI 598
Qy 601 DEEGITRGLIIVRRDWSIAKETQARVLETLKHGDEVEAVRIIVKEVIQKLANYEIPP 660
Db 599 DEEGKVTITRGLIIVRRDWSIAKETQARVLETLKHGDEVEAVRIIVKEVIQKLANYEIPP 658
Qy 661 EKLVIHQITKDSYKAIQPHVAIAKRLAAKGIKVRPGTIISYIVLRSGKTSRDRVILL 720
Db 659 EKLAIVEQITRPLHEYKAIGPHVAVAKKLAAGVKIKPGMVGIVYIVLRGDPISNRILA 718
Qy 721 SEYDPKHKYDPPDYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAML 772
Db 719 EYDPPKHKYDAEYIENQVLPVLRILEFGYRKEDLRYQKTRQVGLTSM 770

RESULT 11
US-07-803-627A-1
; Sequence 1, Application US/07803627A
; Patent No. 5948663
; GENERAL INFORMATION:
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: Purified Thermostable Pyrococcus
; TITLE OF INVENTION: furiosus DNA Polymerase I
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/803,627A
FILING DATE: 02-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/776,552
FILING DATE: 14-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,073
FILING DATE: 19-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,568
FILING DATE: 03-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Barker, M. Paul
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 04121.0004-00
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-803-627A-1

Query Match 78.3%; Score 3151; DB 2; Length 775;
Best Local Similarity 74.6%; Pred. No. 7.4e-237;
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

QY	1	MIFDTYITKDGKPIIRIKPKNGKFKIELDHPQFYIYALLKDSATDEIKAIKGERHG	60
DB	1	MILDVDYITEGKPVIRLFPKNGKFKIEHDTFRFYIYALLRDSKIEEVKKITGERHG	60
QY	61	KIVRVVDVAVKVKKFLGRDVEVVKLIFHPQDVPALRGKIREHPAVIDIYEVDFPFAKY	120
DB	61	KIVRIVDVEKVKKFLGKPIITWKLYLHPQDVPTRKVRHPAVDIFPYDFPFAKY	120
QY	121	LIDKGLIPMEGDEELKLMADFTETHEGDEFGKGEIMISYADBEERAVITWKNIDLPY	180
DB	121	LIDKGLIPMEGDEELKILAFDITLYHEGEEFGKGPIMISYADNEAKVITWKNIDLPY	180
QY	181	VDVNSNEREMIKRFOIVREKDPDVLITVNGNDFLVLKRAEKLGVTLALGRDKEHPE	240
DB	181	VEVVSEREMIKRFLITREKDPDIIVYNGDSFDFPLAKRAEKLGLTKIGRDS--E	238
QY	241	PKIHRMGDSFAVEIKRIFHDFLPVVRRTINLPTVTLAEVAVLGLTKSKLGAEEIAAI	300
DB	239	PKQQRIGDMTAVEVKGRIHFDLYHVTITNLTPTVTLAEVATFGKPKVYADEIAKA	298
QY	301	WETESMKKLAQVSMEDARATVELCKEFPWAEALAKLIGQSVWDVSSTGNLVEWVLL	360
DB	299	WESGENLERVAKYSHEDAKATVELGKFLPMEIQLSRLVGQFLMDVSRSSSTGNLVEWFL	358
QY	361	RVAYERNELAPNPKDEEYRRLRTTYLGGYKPEPGLWENITVLDPRCLYPSIIIVTHN	420
DB	359	KXAYERNEVAPNKPSEEEYQRLRESYTGFGVKEPEKGLWENIVYLDPRALYPSIIITHN	418
QY	421	VSPDTLEREGCKNDVAPIVGVKCKDPGRTIPSLITWREIKKMKATIDPIEKK	480
DB	419	VSPDTLNEGCKNDYDIAFCVGHKFCOIPGIPSLGLHLLEERQKIKTKMKTQDPIEKI	478
QY	481	MLDYQRQAVKLHANSYYGIMGYPKARWYSKECAESVTAWGRHYIEMTIKEEFGPKVL	540
DB	479	LLDYRQKAIKLLANSFYGYGVAKARWYKCAESVTAWGRKYIELWKELEEFPGKVL	538
QY	541	YADTDGFYATPGKEPIYKKAKEFLKYINSKLPGLLELEYEGFYLRGFFVAKRYAVI	600
DB	539	YIDTDGLYATPGGESEIEKKAEFLVYKINSKLPGLLELEYEGFYLRGFFVTKRYAVI	598

QY	601	DEGRITRGLVVRDWESEIAKTOAKVLEAILKEDSVEKAVIYKDVVEEIAKYQVPL	660
DB	599	DEGKVIIRGLVVRDWESEIAKTOARVLEITILKHGDEEAVRIVKEVIQKLANYEIPP	658
QY	661	EKLVIHEQITKDLSEYKAIGHVAIAKRLAAKGIKVRPGTIIISVILRGSGKISDRVILL	720
DB	659	EKLAIYEQITEPLHEYKAIGHVAIAKRLAAKGIKVRPGTIIISVILRGSGKISDRVILL	718
QY	721	SEYDPKHKHDPDYIENQVLPVAVRILEAFGVRKEDLKYQSSKQVGLDAWL	772
DB	719	SEYDPKHKHDAEYIENQVLPVAVRILEGFGVRKEDLKYQTRQVGLTSWL	770

RESULT 12

US-09-244-889A-1
Sequence 1, Application US/09244889A
Patent No. 6489150
GENERAL INFORMATION:
APPLICANT: MATHUR, Eric J.
TITLE OF INVENTION: Purified Thermostable Pyrococcus
furiosus DNA Polymerase I
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunne, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/244,889A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/135,064
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,627
FILING DATE: 02-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/776,552
FILING DATE: 15-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,073
FILING DATE: 19-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,568
FILING DATE: 03-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Barker, M. Paul
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 04121.0004-01
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-244-889A-1

Query Match 78.3%; Score 3151; DB 4; Length 775;
Best Local Similarity 74.6%; Pred. No. 7.4e-237;

Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;
QY 1 MIFDTYITKDGPIIRIFKENGEBKIELEDPHFQVYIVALLKDDSAIDEIKAIKGERHG 60
DB 1 MILDVDYITEGKPVIRLFKENGKFKIEHDTFRFYIYALLRDDSKIEEVKKTIGERHG 60
QY 61 KIVRVDAVKKKFLGRDVEVWKLIFHPQDPVAPALRGKIREHPAVIDIYEDIDIPAKRY 120
DB 61 KIVRIVDVEKVKFLGKPIITVWKLYLEHPQDPVTIREKVRHPAVVDIFEYDIPAKRY 120
QY 121 LIDKGLIPMEGDELKLMADIFETFFVHEGDEFGKEIIMISYADEBEARVITWKNIDLPY 180
DB 121 LIDKGLIPMEGDELKLMADIFETFFVHEGDEFGKEIIMISYADEBEARVITWKNIDLPY 180
QY 181 VDVVSNEREMI KRFVQIVREKDPDLVITNGDNFDLPYLKRAEKLGVTLGRDKEHPE 240
DB 181 VEVVSSSEREMI KRFRIIREKDPDIIVTYNGDSFDFPYLAKRAEKLGIKLTIGRDS--E 238
QY 241 PKIHEMGDSFAVEIKGRIHEDLPVVRTINLPYTLLEAVYEAVLGKTKSGLAEBEIAAI 300
DB 239 PKMQRIGDMTAVEVGRHLHFDLHVITITINLPYTLLEAVYEAVLGKTKSGLAEBEIAAI 300
QY 301 WETESMKLAQYSMEDARATYELGKEFPFMEAEALAKIGQSVMDVSRSTGNLWVWYLL 360
DB 299 WESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVQPLWDSRSTGNLWVWYLL 358
QY 361 RVAYERNELAPNKPDEEYRRLRTYLLGGYVKEPERGLWENITYLDPRCLYPSIIVTHN 420
DB 359 RKAYERNEVAPNKPSEEEYQRLRSYTGFGFVKEPEKGLWENIYLDPRFALYPSIIVTHN 418
QY 421 VSPDTLREGCCKNYDIAPOVGHKFCCKDIPGFIPLSLGHLLERQKIKTKMKETQDPIEKI 478
DB 419 VSPDTLREGCCKNYDIAPOVGHKFCCKDIPGFIPLSLGHLLERQKIKTKMKETQDPIEKI 478
QY 481 MLDYRORAVKLHANSYGYMGYPKARWYKCEASVTANGRHYIEMTKEEKEGFKVL 540
DB 479 LLDYRQAKIULANSFYGYGYAKARYCKEASVTANGRKYIELVWKELEKEGFKVL 538
QY 541 YADTDGFYATIPGKPEIKKAKFEFLKYNKLPGLLELEYEGFYLRGFFVAKRYAVI 600
DB 539 YIDTDGLYATIPGSESEIKKALEFVKYINSKLPGLLELEYEGFYLRGFFVAKRYAVI 598
QY 601 DEGRITRGLVVRDWSSEIAKTQAKLEAILKEDSVKAEIVKDVVEEIAKQVPL 660
DB 599 DEEGKVIPTRGLEIVRWDSEIAKTQARVLETLKHGDVEEAVRIVEIKLANEYIEP 658
QY 661 EKLVIHEOITKDLSEYKALGPVHALAKLAAGIKVRPGTIIIVILRGSGKISDRVILL 720
DB 659 EKLAIYEQITRPLHEYKALGPVAVAKLAAGVYKIEGVMIGVIVLRGDPISNRALLA 718
QY 721 SEYDPKHKYDPPDYIENQVLPVLRILEAFGYRKEDLKYOSSKQVGLDAWL 772
DB 719 BEYDPKHKYDAEYIENQVLPVLRILEGFGYRKEDLRYQKTRQVGLTSLW 770

RESULT 13

US-08-688-649-37

Sequence 37, Application US/08688649

Patent No. 5827716

GENERAL INFORMATION:

APPLICANT: MAMONE, JOSEPH A.

TITLE OF INVENTION: MODIFIED POL-II TYPE DNA

TITLE OF INVENTION: POLYMERASES

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688,649
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wardburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/281
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-688-649-37

Query Match 78.3%; Score 3151; DB 2; Length 776;

Best Local Similarity 74.6%; Pred. No. 7.4e-237;

Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

QY 1 MIFDTYITKDGPIIRIFKENGEBKIELEDPHFQVYIVALLKDDSAIDEIKAIKGERHG 60
DB 1 MILDVDYITEGKPVIRLFKENGKFKIEHDTFRFYIYALLRDDSKIEEVKKTIGERHG 60
QY 61 KIVRVDAVKKKFLGRDVEVWKLIFHPQDPVAPALRGKIREHPAVIDIYEDIDIPAKRY 120
DB 61 KIVRIVDVEKVKFLGKPIITVWKLYLEHPQDPVTIREKVRHPAVVDIFEYDIPAKRY 120
QY 121 LIDKGLIPMEGDELKLMADIFETFFVHEGDEFGKEIIMISYADEBEARVITWKNIDLPY 180
DB 121 LIDKGLIPMEGDELKLMADIFETFFVHEGDEFGKEIIMISYADEBEARVITWKNIDLPY 180
QY 181 VDVVSNEREMI KRFVQIVREKDPDLVITNGDNFDLPYLKRAEKLGVTLGRDKEHPE 240
DB 181 VEVVSSSEREMI KRFRIIREKDPDIIVTYNGDSFDFPYLAKRAEKLGIKLTIGRDS--E 238
QY 241 PKIHEMGDSFAVEIKGRIHEDLPVVRTINLPYTLLEAVYEAVLGKTKSGLAEBEIAAI 300
DB 239 PKMQRIGDMTAVEVGRHLHFDLHVITITINLPYTLLEAVYEAVLGKTKSGLAEBEIAAI 298
QY 301 WETESMKLAQYSMEDARATYELGKEFPFMEAEALAKIGQSVMDVSRSTGNLWVWYLL 360
DB 299 WESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVQPLWDSRSTGNLWVWYLL 358
QY 361 RVAYERNELAPNKPDEEYRRLRTYLLGGYVKEPERGLWENITYLDPRCLYPSIIVTHN 420
DB 359 RKAYERNEVAPNKPSEEEYQRLRSYTGFGFVKEPEKGLWENIYLDPRFALYPSIIVTHN 418
QY 421 VSPDTLREGCCKNYDIAPOVGHKFCCKDIPGFIPLSLGHLLERQKIKTKMKETQDPIEKI 478
DB 419 VSPDTLREGCCKNYDIAPOVGHKFCCKDIPGFIPLSLGHLLERQKIKTKMKETQDPIEKI 478
QY 481 MLDYRORAVKLHANSYGYMGYPKARWYKCEASVTANGRHYIEMTKEEKEGFKVL 540
DB 479 LLDYRQAKIULANSFYGYGYAKARYCKEASVTANGRKYIELVWKELEKEGFKVL 538
QY 541 YADTDGFYATIPGKPEIKKAKFEFLKYNKLPGLLELEYEGFYLRGFFVAKRYAVI 600
DB 539 YIDTDGLYATIPGSESEIKKALEFVKYINSKLPGLLELEYEGFYLRGFFVAKRYAVI 598

Db 1 MILDQVITBKGKGVIRVFKKDGKFKI EYDRFEFEPYIVALLRDDSAIBEIEKITAERHG 60
Qy 61 KIYRVVDVAVKVKKFLGRDVEVWKLIFEHQDVPALR-GKIREHPAVIDIYEYDIPFAKR 119
Db 61 KWKVKRAEKVKKKFLGRSEVWVLYPTHQDVPALRDKIRKHPAVIDIYEYDIPFAKR 120
Qy 120 YLIDKGLIPMEGDEELKMAFDIETFYHGEDEFGKGEIIMISYADEEERARVITWKNIDLP 179
Db 121 YLIDKGLIPMEGDEELKMAFDIETLYHEGEDEFGTGPILMISYADESEARVITWKNIDLP 180
Qy 180 YVDVVSNEREMIKRFVQIVREKDDULITYNGDNFDLPYLIKRAEKLGVTLILGRDKHP 239
Db 181 YVDVVSNEKEMIKRFLKVKVEKOPDVLITYDGNDFPAYLKKRCEKLGVSFTLGRDGS-- 238
Qy 240 EPKIHRRMGDSFAVEIKRIHFDLPFPVVRRTINLPTTYLEAVYEAUAVLGKTKSLGAEIEIAA 299
Db 239 EPKIQRMGDRFAVEKGRHIFDLYPAIRRTINLPTTYLEAVYEAUAVGKPKKVAEIEIAT 298
Qy 300 IWETEESMKLAQVSMEDADATYELGKEFPMAELAKLIGQSVWVDSRSTGNLVWVYL 359
Db 299 AWETGELEGVARYSMEDARVYELGREFFPMEQAQLSRLLGQGLWVDSRSTGNLVWVFL 358
Qy 360 LRVAVERNELAPNKPDEEYRRRLRTTYLGGYVKEPERGLWENITYLDRCLYPSIIVTH 419
Db 359 LRKAYERNELAPNKPDERELARR-RGGYAGYVKEPERGLWENIVYDFSLYPSIIVTH 417
Qy 420 NVSPDTLERGGCKNYDVAPIVGYKFCDFPGFIPSLGELITWROSIKKMKATIDPIEK 479
Db 418 NVSPDTLNRGCKSYDAPQVGHKFCDFPGFIPSLGELNLEERQIKRKMATLDPLER 477
Qy 480 KMLDYRORAVKLHANSYGYMGYPKARWYSKECAESVTAWGRHYEWTIKEIEEKEGFKV 539
Db 478 KLLDYRORAKILANSYGYGYVABARWYCKECAESVTAWGREYIENWIRELEKEGFKD 537
Qy 540 LYADTDGFYATIPGKEPETIKKAKPEFLKYINSKLPGLLELEYEGFYLRGFFVAKKRYAV 599
Db 538 LYADTDGLHATIPGADRETIVKKDLEFLNINPKLPGLLELEYEGFYSGRFFVTKKYAV 597
Qy 600 IDEEGRIITRGLVVRDWSIEIAKTOAKVLEAILKEDSVKAVEIVKQVVEIAKYQVP 659
Db 598 IDEEGKITRGLVVRDWSIEIAKTELARVLEAILRHGDVEAVRIVKEETELSKIEVP 657
Qy 660 LEKLVIEHQITKDSYKAIGPHVAIAKRLAAKGIKVRPCTIISYIVLRGSGKISDRVIL 719
Db 658 PEKLVITEQITRELKDYKATGPHVAIAKRLAARGIKIRPGTVISYIVLKSGRIGDRAIP 717
Qy 720 LSEYDPKHKYDPPDYIENQVLPVLRILEAFGYRKEDLKVQSSKQVGLDAWL 772
Db 718 PDEFDPTKRYDADYIENQVLPVLRILRAFGYKEDERYQKTRQVGLGAWL 770

Search completed: February 9, 2004, 11:53:46

Job time : 34'secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 11:46:55 ; Search time 40 Seconds
(without alignments)
1860.864 Million cell updates/sec

Title: US-09-803-165-34

Perfect score: 4026

Sequence: 1 MIFDTDYITKDGKPIIRIFK.....KEDLKYQSSKQVGLDAMLKK 774

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3356	83.4	1702	2 S42459	DNA-directed DNA p
2	3204.5	79.6	775	2 S67920	DNA-directed DNA p
3	3164.5	78.6	771	2 C75023	DNA polymerase I p
4	3151	78.3	775	2 S35543	DNA-directed DNA p
5	2934.5	72.9	1312	2 S28593	DNA-directed DNA p
6	2886	71.7	1235	2 C1210	probable DNA-direc
7	2767.5	68.7	1670	2 S1551	DNA-directed DNA p
8	1301	32.3	781	2 A69312	DNA polymerase B1
9	1291	32.1	1634	2 E6410	DNA-directed DNA p
10	1209.5	30.0	586	2 C69028	DNA-dependent DNA p
11	1204.5	29.9	781	2 UC7382	DNA-directed DNA p
12	1190	29.6	803	2 B56277	DNA-directed DNA p
13	1138	28.3	784	2 E2515	probable DNA-direc
14	693.5	17.2	1088	2 T05731	DNA-directed DNA p
15	676.5	16.8	768	2 S75407	probable DNA-direc
16	659.5	16.4	1038	1 UC5757	DNA-directed DNA p
17	657.5	16.3	1038	2 T18222	DNA polymerase del
18	657.5	16.3	1086	2 T40242	DNA-directed DNA p
19	649.5	16.1	1086	2 T43266	DNA-directed DNA p
20	626	15.5	901	2 E4210	DNA polymerase B1
21	621.5	15.4	1097	1 RNEYJ3	DNA-directed DNA p
22	619.5	15.4	882	2 S23019	DNA-directed DNA p
23	619.5	15.4	1084	1 S19661	DNA-directed DNA p
24	616.5	15.3	872	2 UC7380	DNA-directed DNA p
25	616.5	15.3	882	2 F90201	DNA polymerase I
26	616.5	15.3	1105	1 S40241	DNA-directed DNA p
27	614.5	15.3	1094	2 S22573	DNA-directed DNA p
28	611	15.2	1107	1 A41618	DNA-directed DNA p
29	609	15.1	1513	2 T28158	probable DNA-direc

30	605.5	15.0	1106	1 A39299	DNA-directed DNA p
31	587	14.6	1492	2 T18560	DNA-directed DNA p
32	584.5	14.5	879	2 A56277	DNA-directed DNA p
33	583.5	14.5	1462	1 DJHUAC	hypothetical prote
34	581	14.4	1081	2 T20698	DNA-directed DNA p
35	578.5	14.4	875	2 JCS186	probable DNA-direc
36	570.5	14.2	959	2 F72763	DNA-directed DNA p
37	570	14.2	1465	2 S45628	DNA-dependent DNA p
38	559.5	13.9	223	2 E69125	DNA-directed DNA p
39	548.5	13.6	1339	1 S20052	DNA-directed DNA p
40	527.5	13.1	1468	2 S58250	DNA-directed DNA p
41	522.5	13.0	1505	2 S28079	DNA-directed DNA p
42	516.5	12.8	844	2 T31321	DNA-directed DNA p
43	516	12.8	1405	1 DJZPA	hypothetical prote
44	504.5	12.5	2154	2 A84669	DNA-directed DNA p
45	500.5	12.4	913	2 T17675	DNA-directed DNA p

ALIGNMENTS

RESULT 1

S42459

DNA-directed DNA polymerase (EC 2.7.7.7) Vent, intein containing precursor - Thermococcc
N;Contains: DNA endonuclease (EC 3.1.1.-) PI-I; DNA endonuclease (EC 3.1.1.-) PI-II; DN
C;Species: Thermococcus litoralis

C;Date: 31-Dec-1993 #sequence revision 02-Aug-1994 #text_change 18-Jun-1999

C;Accession: S42459; S42451; S42450; S42458

R;Perler, F.B.; Comb, D.G.; Jack, W.E.; Moran, L.S.; Qiang, B.; Kucera, R.B.; Benner, J
submitted to the EMBL Data Library, September 1992

A;Reference number: S42458

A;Accession: S42459

A;Molecule type: DNA

A;Residues: 1-1702 <PER1>

A;Cross-references: EMBL:M74198; NID:G154685; PIDN:AA72100.1; PID:G154686

R;Perler, F.B.; Comb, D.G.; Jack, W.E.; Moran, L.S.; Qiang, B.; Kucera, R.B.; Benner, J
Proc. Natl. Acad. Sci. U.S.A. 89, 5577-5581, 1992

A;Title: Intervening sequences in an Archaea DNA polymerase gene.

A;Reference number: S42450; MUID:92302285; PMID:1608969

A;Accession: S42451

A;Molecule type: DNA

A;Residues: 181-222;387-425;452-476;483-524;1021-1062;1076-1099;1466-1489;1533-1547 <PE

A;Cross-references: EMBL:M74198

R;Hodges, R.A.; Perler, F.B.; Noren, C.J.; Jack, W.E.

Nucleic Acids Res. 20, 6153-6157, 1992

A;Title: Protein splicing removes intervening sequences in an archaea DNA polymerase.

A;Reference number: S40786; MUID:93117083; PMID:1475179

C;Contents: annotation

C;Function: <VENT>

A;Description: nucleotidyltransferase

A;Note: DNA-directed DNA polymerase Vent

C;Function: <END1>

A;Description: endonuclease; hydrolase

A;Note: DNA endonuclease PI-Tlii

C;Function: <END2>

A;Description: endonuclease; hydrolase

A;Note: DNA endonuclease PI-Tlii

C;Superfamily: Thermococcus litoralis DNA-directed DNA polymerase Vent

C;Keywords: DNA binding; endonuclease; hydrolase; nucleotidyltransferase; protein splic

F;1-494;1033-1081;1472-1702/Product; DNA-directed DNA polymerase Vent #status predicted

F;1-494/DNA; DNA-directed DNA polymerase Vent extein 1 #status predicted <XT1>

F;495-1032/Product; DNA endonuclease PI-II (pol Vent intein 1) #status predicted <MAT2>

F;1033-1081/DNA; DNA-directed DNA polymerase Vent extein 2 #status predicted <XT2>

F;1082-1471/Product; DNA endonuclease PI-I (pol Vent intein 2) #status predicted <XT2>

F;1472-1702/DNA; DNA-directed DNA polymerase Vent extein 3 #status predicted <MAT3>

F;494-1033/Cross-link; peptide (Asn-Ser) #status predicted

F;1081-1472/Cross-link; peptide (Asp-Thr) #status predicted

Query Match 83.4%; Score 3356; DB 2; Length 1702;

Best Local Similarity 42.7%; Pred. No. 2e-164;

Matches 726; Conservative 33; Mismatches 15; Indels 928; Gaps 2;

QY 1 MIFDTDYITKDGKPIIRIFKENGFKIELDPHFQPIYALLKDDSAIDKAIKGRHG 60

Db 1 MILDVDITKGGKPIIRIFKENGEPKIELDPHFQPIYALLKDDSAIEBIKAIGERHG 60
QY 61 KIVRVDAVKKKFLGRDVEVWKLIFEHQDDPALRGKIREHPAVDIDYEYDIPAKY 120
Db 61 KTVRLDAVKKKFLGRDVEVWKLIFEHQDDPALRGKIREHPAVDIDYEYDIPAKY 120
QY 121 LIDKGLIPMEGDEBELKMAFDIETFYHEGDEFKGEIIMISYADEEAEARVITWKNIDL 180
Db 121 LIDKGLIPMEGDEBELKMAFDIETFYHEGDEFKGEIIMISYADEEAEARVITWKNIDL 180
QY 181 VDVVSNREMIKRFVQVREKDPVLITYNGDNFDLPYLKRAEKLGVTLILGSDKEHPE 240
Db 181 VDVVSNREMIKRFVQVREKDPVLITYNGDNFDLPYLKRAEKLGVTLILGSDKEHPE 240
QY 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTLAEVAVLGTGKSLGAEEIAAI 300
Db 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTLAEVAVLGTGKSLGAEEIAAI 300
QY 301 WETESMKKLAQYMEDARATYELGKFFPMEASLAKLIGOSVWDSRSSTGNLVEYLL 360
Db 301 WETESMKKLAQYMEDARATYELGKFFPMEASLAKLIGOSVWDSRSSTGNLVEYLL 360
QY 361 RVAVARNELAPNKDEEYRRRLATTYLGGVKPEPERGLMENITYLDFRCLYPSIIVTHN 420
Db 361 RVAVARNELAPNKDEEYRRRLATTYLGGVKPEPERGLMENITYLDFRCLYPSIIVTHN 420
QY 421 VSPDTLEBEGKNVDVPIGVKCPKPPGPIPSILGELITMROEIKKKAKAIIDPIEKK 480
Db 421 VSPDTLEBEGKNVDVPIGVKCPKPPGPIPSILGELITMROEIKKKAKAIIDPIEKK 480
QY 481 MLDVQRQAVKLAH----- 493
Db 481 MLDVQRQAVKLAH----- 493
QY 494 ----- 493
Db 541 VNNLFAFSFNKKIKSEVKKVKKALIRHKYKGAVEIQLSGGRKINITAGHSFTVRNGEI 600
QY 494 ----- 493
Db 601 KEVSGGKIEGDLIVAPKKIKLNEKGVSNIPELISDSEETADIVMTISAKGRKFFK 660
QY 494 ----- 493
Db 661 GMLRLRMFGENRRIRTFNRYLPHLEKGLIKLLPRGYEVDWERLKKYQLYEKL 720
QY 494 ----- 493
Db 721 SVKYNKREYLVMPNEIKDFISYPPQKELEWKTGTNGFRNCILKVDDEDFGKLIGY 780
QY 494 ----- 493
Db 781 VSEGVAGAKNKTGGISYVLYNEDPNVLESKMKNAEKKFFGKVRVDRNCVSIKGMAYL 840
QY 494 ----- 493
Db 841 VMKCLCGALAEKRIPIVILTSPEPVRSFLEAYTGGDTHPSKRFSLTSKELLANQL 900
QY 494 ----- 493
Db 901 VFLNLSLGISVKGIFGDSGVYRVYINEDLQFPQTSREKNTYVSNLIPKEILRDVFGKEF 960
QY 494 ----- 493
Db 961 KMTFPKKFELVDGKLNREKAKLEFFINGDILVDRVKSVEKDEYGVYVVDLSVDENEN 1020
QY 494 -----NSYCYMGYPKARWYSKECAESVTAWGRHYIEMTIKIEBKFGFKVLYA 542
Db 1021 FLVGFGLLYAHNSYCYMGYPKARWYSKECAESVTAWGRHYIEMTIKIEBKFGFKVLYA 1080
QY 543 D----- 543

Db 1081 DSVGSESEIIIRONGKIRFVKIKOLFQSKVDYSIGKEYCILEGVEALTLDGDKLVWKP 1140
QY 544 ----- 543
Db 1141 PYVMEHRANKBMFRILWNTSNWIDVTEHDSLIGYLNSTKTKTAKKIGERLKEVPFELGK 1200
QY 544 ----- 543
Db 1201 AVKSLICPNAPLKQENTKTSIAVFWELVGLVGNWGGDSRWAEYVYLGSLGKDAABE 1260
QY 544 ----- 543
Db 1261 IKQKLELPLKTYGVNISVYYPKNEKGFNLLAKSLVKFMKGRHFKDEKGRKIPEFMYELPV 1320
QY 544 ----- 543
Db 1321 TYIEAFLRGLFSADGTVTIRKGVPEIRLTNIDADFLEVRKLLWIVGISNISIFAETTPNR 1380
QY 544 ----- 543
Db 1381 YNGVSTGYYSKHLRIKRWFAERIGFLLERKQELLEHLKSARVKRNTIDFGPDLVHVK 1440
QY 544 -----TDGFYATIPGEKPEETIKKAKKEFLKYINS 572
Db 1441 KVEEIPYEGYVYDIEVEETHRFFANNILVHNTDGFYATIPGEKPEELIKKAKKEFLKYINS 1500
QY 573 KLPGLLELEYEGFYLRGFFVAKRVAVIDBERGRIITRGLVVRDWSIAKETQAKVLEA 632
Db 1501 KLPGLLELEYEGFYLRGFFVAKRVAVIDBERGRIITRGLVVRDWSIAKETQAKVLEA 1560
QY 633 ILKESDVEKAVEIKVDVVEEIAKYQVPLEKVIHEQITKQJSEYKAIQPHVAIAKRLAAK 692
Db 1561 ILKESDVEKAVEIKVDVVEEIAKYQVPLEKVIHEQITKQJSEYKAIQPHVAIAKRLAAK 1620
QY 693 GKWPRTIISVILRGSKISDRVILLSEVDPKXKHYDDPYVITENOVLPVLRILEAFG 752
Db 1621 GKWPRTIISVILRGSKISDRVILLSEVDPKXKHYDDPYVITENOVLPVLRILEAFG 1680
QY 753 YRKEDLYOSSKQVGLDAWLKK 774
Db 1681 YRKEDLYOSSKQVGLDAWLKK 1702
RESULT 2
S67920
DNA-directed DNA polymerase (EC 2.7.7.7) - Thermococcus sp.
C;Species: Thermococcus sp.
C;Date: 17-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 18-Jun-1999
C;Accession: S67920
R;Southworth, M.S.; Kong, H.; Kucera, R.B.; Ware, J.; Jannasch, H.W.; Perler, F.B.
submitted to the EMBL Data Library, January 1996
A;Description: Cloning, expression and modulation of the 3'-5' exonuclease activity of
A;Reference number: S67920
A;Accession: S67920
A;Molecule type: DNA
A;Residues: 1-775 <SOU>
A;Cross-references: GB:U47108; NID:gl197451; PIDN:AAA88769.1; PID:gl197452
A;Experimental source: strain 90N-7
C;Superfamily: Herpesvirus DNA-directed DNA polymerase
C;Keywords: DNA binding; nucleotidyltransferase
Query Match 79.6%; Score 3204.5; DB 2; Length 775;
Best Local Similarity 76.6%; Pred. No. 4.1e-157;
Matches 592; Conservative 89; Mismatches 89; Indels 3; Gaps 2;
QY 1 MIPDTYITKGGKPIIRIFKENGEPKIELDPHFQPIYALLKDDSAIEBIKAIGERHG 60
Db 1 MILDVDITKGGKPIIRIFKENGEPKIELDPHFQPIYALLKDDSAIEBIKAIGERHG 60
QY 61 KIVRVDAVKKKFLGRDVEVWKLIFEHQDDPALRGKIREHPAVDIDYEYDIPAKY 120
Db 61 KTVRLDAVKKKFLGRDVEVWKLIFEHQDDPALRGKIREHPAVDIDYEYDIPAKY 120
QY 121 LIDKGLIPMEGDEBELKMAFDIETFYHEGDEFKGEIIMISYADEEAEARVITWKNIDL 180
Db 121 LIDKGLIPMEGDEBELKMAFDIETFYHEGDEFKGEIIMISYADEEAEARVITWKNIDL 180
QY 181 VDVVSNREMIKRFVQVREKDPVLITYNGDNFDLPYLKRAEKLGVTLILGSDKEHPE 240
Db 181 VDVVSNREMIKRFVQVREKDPVLITYNGDNFDLPYLKRAEKLGVTLILGSDKEHPE 240
QY 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTLAEVAVLGTGKSLGAEEIAAI 300
Db 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTLAEVAVLGTGKSLGAEEIAAI 300
QY 301 WETESMKKLAQYMEDARATYELGKFFPMEASLAKLIGOSVWDSRSSTGNLVEYLL 360
Db 301 WETESMKKLAQYMEDARATYELGKFFPMEASLAKLIGOSVWDSRSSTGNLVEYLL 360
QY 361 RVAVARNELAPNKDEEYRRRLATTYLGGVKPEPERGLMENITYLDFRCLYPSIIVTHN 420
Db 361 RVAVARNELAPNKDEEYRRRLATTYLGGVKPEPERGLMENITYLDFRCLYPSIIVTHN 420
QY 421 VSPDTLEBEGKNVDVPIGVKCPKPPGPIPSILGELITMROEIKKKAKAIIDPIEKK 480
Db 421 VSPDTLEBEGKNVDVPIGVKCPKPPGPIPSILGELITMROEIKKKAKAIIDPIEKK 480
QY 481 MLDVQRQAVKLAH----- 493
Db 481 MLDVQRQAVKLAH----- 493
QY 494 ----- 493
Db 541 VNNLFAFSFNKKIKSEVKKVKKALIRHKYKGAVEIQLSGGRKINITAGHSFTVRNGEI 600
QY 494 ----- 493
Db 601 KEVSGGKIEGDLIVAPKKIKLNEKGVSNIPELISDSEETADIVMTISAKGRKFFK 660
QY 494 ----- 493
Db 661 GMLRLRMFGENRRIRTFNRYLPHLEKGLIKLLPRGYEVDWERLKKYQLYEKL 720
QY 494 ----- 493
Db 721 SVKYNKREYLVMPNEIKDFISYPPQKELEWKTGTNGFRNCILKVDDEDFGKLIGY 780
QY 494 ----- 493
Db 781 VSEGVAGAKNKTGGISYVLYNEDPNVLESKMKNAEKKFFGKVRVDRNCVSIKGMAYL 840
QY 494 ----- 493
Db 841 VMKCLCGALAEKRIPIVILTSPEPVRSFLEAYTGGDTHPSKRFSLTSKELLANQL 900
QY 494 ----- 493
Db 901 VFLNLSLGISVKGIFGDSGVYRVYINEDLQFPQTSREKNTYVSNLIPKEILRDVFGKEF 960
QY 494 ----- 493
Db 961 KMTFPKKFELVDGKLNREKAKLEFFINGDILVDRVKSVEKDEYGVYVVDLSVDENEN 1020
QY 494 -----NSYCYMGYPKARWYSKECAESVTAWGRHYIEMTIKIEBKFGFKVLYA 542
Db 1021 FLVGFGLLYAHNSYCYMGYPKARWYSKECAESVTAWGRHYIEMTIKIEBKFGFKVLYA 1080
QY 543 D----- 543

QY	121	LIDKGLIPMGDEBELKUMAFDITFTYHEGDEFGKEIIMISYADEEFARVITWKNIDL	180
Db	121	LIDKGLIPMGDEBELKUMAFDITFTYHEGDEFGKEIIMISYADEEFARVITWKNIDL	180
QY	181	VDVYSNEREMIKFVQIVREKDPDLVLTITNGDNFDPYLKIKRAEKLGVTLGLGRDKEHPE	240
Db	181	VDVVSTEKEMIKFLRVREKDPDLVLTITNGDNFDPYLKIKRAEKLGVTLGLGRDKEHPE	240
QY	241	PKTHRMGDSFAVITKGRIHFDLPVVRRTTNLTPTTLEAVYAVLGTKSKLGAETIAI	300
Db	239	PKIQRMGDRFAVEVKGHFDLPVIVRTTNLTPTTLEAVYAVLGTKSKLGAETIAI	298
QY	301	WETEESMKLAQYSMEDARATYELGKEFFPWEAEALAKLIGOSVMDVSRSSTGNLVEWYLL	360
Db	299	WESGEGLEVARYSMEDAKVYELGREFFPWEAQLSRLIGOSLWDVSRSSTGNLVEWFL	358
QY	361	RVAYERNELAPNKPDEEYRRRLRTTTLVGYVKEPERGLWENTTYLDFRCLYPSIIVTHN	420
Db	359	RKAYRNELAPNKPDEELARR-RGGVAGYVKEPERGLWMDNIVYLDLFRSLYPSIIITHN	417
QY	421	VSPDTLRECKKNYDVAPIVGXYCKDFPGFIPISILGELITMRQEIKKKKATIDPIEKK	480
Db	418	VSPDTLNRCKEYDVAVEVGHKECKDFPGFIPISLLGLDLEERQKIKRKKAVIDPLEKK	477
QY	481	MLDIYRQBAVKLHANSYGYGYPKARYWSECAESVTAWGRHYIEMTIKEIEKPGFKVL	540
Db	478	LLDIYRQBAIKILANSFYGYGYAKARYWSECAESVTAWGREYIEMVIRELEKPGFKVL	537
QY	541	YADTDGFYATIPGKEKPTIKKKAKEFLKYINSKLPGLLELEYEGFFYLRGFFVAKKRYAVI	600
Db	538	YADTDGLHATIPGADAETVKKKAKEFLKYINPKLPGLLELEYEGFFYLRGFFVAKKRYAVI	597
QY	601	DEGRITTRGLEVVRRDWSIAKETQAKVLEAILKEDSVEKAVEIKVDVVEEIAKYQVPL	660
Db	598	DEGKITTRGLEIVRRDWSIAKETQARVLEAILKHGDVEEAVRIVEVTEKLSKYEVP	657
QY	661	EKLVIHQITKOLSEYKAIGPHVAIAKRLAAKGKVRPGTIIISVILRGSGKISDRVILL	720
Db	658	EKLVIHQITRLRDYKATGPHVAVAKRLAARGVKIRPGTIVISYILRGSGRIGDRAIPA	717
QY	721	SEVDPKHKHYDDPYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAMLK	773
Db	718	DEFDPTKHYDAEYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAMLK	770

RESULT 3
 C75023
 DNA polymerase I PAB1128 - Pyrococcus abyssi (strain Orsay)
 C/Species: Pyrococcus abyssi
 C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C/Accession: C75023
 R;anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str
 A;Reference number: A75001
 A;Accession: C75023
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-771 <RAW>
 A;Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PID:g5458960; PID:g545913
 A;Experimental source: strain Orsay
 C;Genetics:
 A;Gene: polI; PAB1128
 C;Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match	78.6%;	Score 3164.5;	DB 2;	Length 771;
Best Local Similarity	76.1%;	Pred. No. 4.6e-155;		
Matches 588;	Conservative 89;	Mismatches 93;	Indels 3;	Gaps 2;

QY	1	MIEDTYITKDGPIIRIFKNGEKFIEDPHQPIYIALLKDDSAIDBIKAGHERG	60
Db	1	MIIDADYITEDGPIIRIFKNGEKFVEDTRFRPIYIALLKDDSAIDBIKAGHERG	60

QY 61 KIVRVDAVAVKKKFLGRDVEWVKLI FEHPQDVDPALRKIREHPAVIDIYEYDIPFAKRY 120
DB : : : : :
QY 61 KIVRI TEVEKVQKFKLGRPIEVWKLYLEHFDQVPALREKIREHPAVDIFEYDIPFAKRY 120
DB : : : : :
QY 121 LIDKGLIPMEGDRELKLMAFDITTFHEGDEFKGGEIMISYADEEBARVITWNKIDLPY 180
DB : : : : :
QY 121 LIDKGLTPMGNBELTFLAVDIETLHYEGBEFGKGPIMISYADEEAGKVLITWKSIDLUPY 180
DB : : : : :
QY 181 VDVYSNEREMIKFVOIVREKDDPVATITNGNDPDLPLYLKRAEKLGVTLLGLGRDKHEPE 240
DB : : : : :
QY 181 VEVSSEEREMI KLKVKVIREKD PDVIITIYNGDNDFPFYLLKRAEKLGIKPLGRD -NSE 238
DB : : : : :
QY 241 PKTHRWGDSFAVBIKGRIHFDLPFPVVRTINLTPTYLEAVYEAVALGTKSKLGAEEATAAI 300
DB : : : : :
QY 239 PKQWRMGDSLAVBIKGRIHFDLPFPVVRTINLTPTYLEAVYEAIFGSKKEKYAHEAEA 298
DB : : : : :
QY 301 WEYESNMKKLAQSMEDARATYELGCEFFPWAEALAKIGOSVMDVSRSGTNLVWVYL 360
DB : : : : :
QY 299 WETGKGLERVAKYSMEDAKVTFELGKEFFPWAEQLARLVGQPWWDVSRSGTNLVWFLL 358
DB : : : : :
QY 361 RVAYERNELAPNKPDEBEEYRRRLRTTYLGGYKXEPERGLWENITYLDFRCLYPSTIITHN 420
DB : : : : :
QY 359 RKAYERNELAPNKPDEREYERRLRSEYEGGYVKEPEKGLMEGVSLDFRSLYPSIIITHN 418
DB : : : : :
QY 421 VSPDTLEREGCKNVDAPIVGYYKCFDPPGFIPSIILGELITMTROEIKKOMKATIDPIBK 480
DB : : : : :
QY 419 VSPDTLNRENCKEYDVAPOVGHFRCDKDFPGFIPLNLGILLEEROQIKKRKESKDQPEKK 478
DB : : : : :
QY 481 MLDYORAVKLVHANSYGYNGYKPARWYSKECAESVTAWGRHYIEMTKIEEEKFGPKVL 540
DB : : : : :
QY 479 LLDYORAIKI LANSYGYGYGAKARWYKCECAESVTAWGRQYIDLVRERELES- GPKVL 537
DB : : : : :
QY 541 YADTDGFYATI PGEKPETIKKAKEFLKYINSKLPGLLELEYEGFYLRGFTVAKKRYAVI 600
DB : : : : :
QY 538 YITDGLYATIPAKHEEIKEKALKFVEYINSKLPGLLELEYEGFYARGFTVTKKRYALI 597
DB : : : : :
QY 601 DEGRITTRGLEVVVRDWSIAKETQAKVLEAILKEDSVEKAIVEIKDWVEEIAKYQVPL 660
DB : : : : :
QY 598 DERGKI VTRGLEIVRDWSIAKETQAKVLEAILKHGNVD EAVKIVKEVTEKLSKYEIPP 657
DB : : : : :
QY 661 EKLVIEHQITKDISYKATGPHVAIAKRIAAKGIKYPGTII SVILRGSGKISRIVILL 720
DB : : : : :
QY 658 EKUVIIEQITRPUSEYKATGPHVAIRKAAKGVKVPGMVIGYIVLRGDPGISKRAIAI 717
DB : : : : :
QY 721 SEYDPKHKHXDPDYIENQVLPVLRILEAFGYRKEDLKYQSSKQVGLDAWLK 773
DB : : : : :
QY 718 EEFDPKHKXYDAEYIENQVLPVLRILRAFGYRKEDLKYQTKQVGLGAWLK 770
DB : : : : :

RESULT 4

S35543

DNA-directed DNA polymerase (EC 2.7.7) - Pyrococcus furiosus

C;Species: Pyrococcus furiosus

C;Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 20-Jun-2000

C;Accession: S35543; S44596

R;Umori, T.; Ishino, Y.; Toh, H.; Asada, K.; Kato, I.

Nucleic Acids Res. 21, 259-265, 1993

A;Title: Organization and nucleotide sequence of the DNA polymerase gene from the archaeal thermophilic bacterium Pyrococcus furiosus

A;Reference number: S35543; MUID:93181200; PMID:8441634

A;Accession: S35543

A;Molecule type: DNA

A;Residues: 1-775 <UM1>

A;Cross-references: EMBL:D12983; NID:g216917; PIDN:BAA02362.1; PID:g216918

A;Accession: S44596

A;Molecule type: protein

A;Residues: 1-12 <UM2>

C;Genetics:

A;Gene: pol

C;Superfamily: herpesvirus DNA-directed DNA polymerase

C;Keywords: DNA binding; nucleotidyltransferase

F;1-775/Product: DNA-directed DNA polymerase #status experimental <MAT>

Query Match 78.3%; Score 3151; DB 2; Length 775;

```

Query Match
78.3%; Score 3151; DB 2; Length 775;

```


Best Local Similarity 74.6%; Pred. No. 2,3e-154;
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

```
QY 1 MFDYDITKDGKPIIRIFKNGEFGKIELDPHFQPIYIALLKDDSAIDBIKAIKGRHG 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MILDVEDYITEGKPIIRIFKNGEFGKIELDPHFQPIYIALLKDDSAIDBIKAIKGRHG 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 KIVRVVDVAVKVKKFLGRDVEVWKLIFEPHQPDPALRGKIREHPAVDIYEDIPFAKRY 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 KIVRIVDVAVKVKKFLGRDVEVWKLIFEPHQPDPALRGKIREHPAVDIYEDIPFAKRY 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 LIDKGLIPMEGDEBELKMAFDIETFYHGDGDFGKELIIMISYADEEERARVITKNIDLPY 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LIDKGLIPMEGDEBELKMAFDIETFYHGDGDFGKELIIMISYADEEERARVITKNIDLPY 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 VDVVSNEREMIKFVQIVREKDPDLITVNGDNFDPYIKAEKLGVTLLGRDKHEPE 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 VDVVSNEREMIKFVQIVREKDPDLITVNGDNFDPYIKAEKLGVTLLGRDKHEPE 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTLVAVYVAVLQKTKSGABEIAI 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTLVAVYVAVLQKTKSGABEIAI 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 299 PKMQRIGDMTAVEKGRHFDLPYTLVAVYVAVLQKTKSGABEIAI 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 PKMQRIGDMTAVEKGRHFDLPYTLVAVYVAVLQKTKSGABEIAI 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 RVAYERNELAPNKPDEEYRRRLRTYLGKYGKPERGLWENITVLDPRCLYPSIIIVTHN 420
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 RVAYERNELAPNKPDEEYRRRLRTYLGKYGKPERGLWENITVLDPRCLYPSIIIVTHN 420
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 359 RKAYERNELAPNKPDEREYRRRLRESYAGGVYKPEKGLWGLVSLDFRSYPSIIIVTHN 418
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 RKAYERNELAPNKPDEREYRRRLRESYAGGVYKPEKGLWGLVSLDFRSYPSIIIVTHN 418
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 VSPDTLREGCKNVDVAVIVGVYKFCPPGPIPSILGELITWROEIKKMKATIDPIEKK 480
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 VSPDTLREGCKNVDVAVIVGVYKFCPPGPIPSILGELITWROEIKKMKATIDPIEKK 480
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 419 VSPDTLREGCKNVDVAVIVGVYKFCPPGPIPSILGELITWROEIKKMKATIDPIEKK 478
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 VSPDTLREGCKNVDVAVIVGVYKFCPPGPIPSILGELITWROEIKKMKATIDPIEKK 478
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 481 MLDYQRAVK----- 490
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 MLDYQRAVK----- 490
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 479 MLDYQRAVKILANSILPEWVPLIKNGKVIKIFRIGDFVGLMKANQKVKTKGTGVLE 538
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 479 MLDYQRAVKILANSILPEWVPLIKNGKVIKIFRIGDFVGLMKANQKVKTKGTGVLE 538
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 491 ---LHA----- 493
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 491 ---LHA----- 493
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 539 VAGIHAFSFDKSKKARVMAVKAVIRHRYSGNVYIRVNLNLSGRKITTTEGHSFLVYRNGDL 598
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 539 VAGIHAFSFDKSKKARVMAVKAVIRHRYSGNVYIRVNLNLSGRKITTTEGHSFLVYRNGDL 598
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 494 ----- 493
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 494 ----- 493
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 599 VEATGEDVKIGDILAVFRSVNLPEKEERLNIIVELLNLNLSPEETEDIILTIPIVKGKKNFPK 658
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 599 VEATGEDVKIGDILAVFRSVNLPEKEERLNIIVELLNLNLSPEETEDIILTIPIVKGKKNFPK 658
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 494 ----- 493
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 494 ----- 493
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 659 GMLRTLRWIFGEKRVRTASRYLRHLNLYIRLRKIGYDIIIDKSGLEKRYLYEKLVDV 718
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 659 GMLRTLRWIFGEKRVRTASRYLRHLNLYIRLRKIGYDIIIDKSGLEKRYLYEKLVDV 718
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 5

S68593
DNA-directed DNA polymerase (EC 2.7.7.7) Vent, intein containing precursor - Pyrococcus
N:Contains: DNA endonuclease (EC 3.1.-.-) PI-I; DNA-directed DNA polymerase (EC 2.7.7.7)
C:Species: Pyrococcus sp.
C:Date: 24-Aug-1996 #sequence_revision 01-Nov-1996 #text_change 24-Sep-1999
C:Accession: S68593
R:Xu, M.Q.; Southworth, M.W.; Merzha, F.B.; Hornstra, L.J.; Perler, F.B.
submitted to the EMBL Data Library, August 1993
A:Description: In vitro protein splicing of purified precursor and the identification of
A:Reference number: S68593
A:Accession: S68593
A:Molecule type: DNA
A:Residues: 1-1312 <XUA>
A:Cross-references: EMBL:U00707; NID:G436492; PDB:AAA67130.1; PID:9825735
R:Xu, M.Q.; Southworth, M.W.; Merzha, F.B.; Hornstra, L.J.; Perler, F.B.
Cell 75, 1371-1377, 1993

```
QY 494 ----- 493
Db 719 VRNGNKRVLVEFNAVRDVISLMPBELKEMWIGTRNGFRMGTFVDIDEDFAKLLGYVY 778
QY 494 ----- 493
Db 779 SEGSAKWKQNTGGWSYTVLNYNENDEVLDDMEHLAKFPFGKVGKGVVPEIPKWAYII 838
QY 494 ----- 493
Db 839 PESLGTALBNKRVPDEVFTSSKGVWAFLEGYFIGDGDVHPKRVRLSTKSELLVNGIV 898
QY 494 ----- 493
Db 899 LLLNSLGSALILGDSGVVYVYVNEELKFTFYRKKKNVYHSHVIPKDLIKETFGKVPQK 958
QY 494 ----- 493
Db 959 NISYKKFRELIVNGKLDREKAKRIEALLNGDIVLDRWBEIKREYDYVYDLSVDEDNF 1018
QY 494 ----- 493
Db 1019 LAGFGFLYAHNSYGYGYGAKARWYCKEACSVTAWGREYIEBFVRKELEKEGFKVLID 1078
QY 544 TQCFYATIPGEPETIKKAKFPLKYNSKLPGLLELEYEGPYLRGFFVAKRYAVIDEE 603
Db 1079 TDGLYATIPGAKPEEIKKALEFVDYINAKLPGLLELEYEGPYVRGFFVTKKYYALIDEE 1138
QY 604 GRITTRGLEVWRDSEIAKETQAKVLRAILKEDSVKAVEIVKDVVEIAKYQVPLEKL 663
Db 1139 GXIITRGLIVRDSSEIAKETQAKVLRAILKGVNEAVKIVKEVTEKLSYIEPPKLL 1198
QY 664 VIHQITKDSLEYKAGHVAIAKRAAGIKVRPCTIISYIVLRGSGKISDRVILLSEY 723
Db 1199 VYIEQITRELPHEYKAGHPHVAIAKRAAGIKVRPCTIISYIVLRGSGKISDRVILLSEY 1258
QY 724 DPKKXKYDPDYIENQVLPVLRILEAFGRYKEDLYQSSKQVGLDNL 772
Db 1259 DLKHKYDAEYIENQVLPVLRILEAFGRYKEDLYQSSKQVGLDNL 1307

RESULT 6
C71210
probable DNA-directed DNA polymerase - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: C71210
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: C71210
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1235 <KAW>
A:Cross-references: GB:AP000007; NID:G3236134; PIDN:BAA31074.1; PID:G3258391
A:Experimental source: strain OT3
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1947
C:Superfamily: hypothetical protein PH0202

Query Match 71.7%; Score 2886; DB 2; Length 1235;
Best Local Similarity 46.5%; Pred. No. 1.7e-140;
Matches 574; Conservative 107; Mismatches 89; Indels 464; Gaps 5;

QY 1 MIFDTDTYTKGKPIIRIFKVENGEFKLELDPHQPIYVYALLKDDSDAIDEIKAIKGRHG 60
Db 1 MLDADYITDGGKPIIRIFKVENGEFKLELDPHQPIYVYALLKDDSDAIDEIKAIKGRHG 60
QY 61 KIVRVVDVAVKKKFLGRDVEVWKLIFEPHPODVPAIRCKIREHPAVIDIYEYDIPFAKRY 120
```

```
Db 61 KVAIVATEKIQRFGLGRPIEVWKLILEHPQDVPAIRDKIREHPAVVDIYEYDIPFAKRY 120
QY 121 LIDKGLIPMEGDGEBELKMAPDIETFYHEGDFGGEIIMISYADEERARVITWKNIDLPY 180
Db 121 LIDKGLIPMEGNEKLTFLAVDITELYHEGDFGGEIIMISYADEERARVITWKNIDLPY 180
QY 181 VDVYSNEREMIKRQVQVIREKDPDLITYNGDNPDLPYLKRAEKLGVTLILLGRDKEHPE 240
Db 181 VEVVYSNEREMIKRQVQVIREKDPDLITYNGDNPDLPYLKRAEKLGVTLILLGRD--NSE 238
QY 241 PKIHRMGDSFAVEIKGRIHFDLPVVRRTINLPYTLLEAVYEAIVGKTCKLGAEEITAAI 300
Db 239 PKQKMGDSLAVEIKGRIHFDLPVVRRTINLPYTLLEAVYEAIVGKTCKLGAEEITAAI 298
QY 301 WETESWKKLAQVSMEDARATYELGKFPFPEAEALAKLQCSVMDVSRSTGNLVEWYLL 360
Db 299 WETGEGELVAKYSMEDAKVTYELGKFPFPEAEALAKLQCSVMDVSRSTGNLVEWYLL 358
QY 361 RVAYERNELAPNKPDEEYERRLRTTVLGGYKPEPGLWENITYLDLFRCLYPSIIVTHN 420
Db 359 RYAYERNELAPNKPDEEYERRLRTTVLGGYKPEPGLWENITYLDLFRCLYPSIIVTHN 418
QY 421 VSPDTLREGCKNYDVAPIVGYKFCDFPGFIPSIILGELITMRQEIKKMKATIDPIEKK 480
Db 419 VSPDTLNREGCEYDVAPKVGHRFCDFPGFIPSIILGELITMRQEIKKMKESKDPVEKK 478
QY 481 MLDYRQBAVKL----- 491
Db 479 LLDYRQBAVKL----- 491
QY 492 ----- 491
Db 539 VKDLKALSPNRETCKSELKVKALIRHRYSGKYSIKLSGRRIKITSGHLSFVKNGL 598
QY 492 ----- 491
Db 599 VKRGDELKPGDLVVVPGRLKLPESQVNLVELLLKPEEETSNVIMIPVAKRKNFK 658
QY 492 ----- 491
Db 659 GMLKTLWIFGEGEPRTAGRYKHLERLGYVKLKERGCEVLDWESLKRKYKLYETLKN 718
QY 492 ----- 491
Db 719 LKYNNGSRAYMVEFNLSRDVSLMPEELKEWIIIGEPRGPKIGTFIDVDDSPAKLLGYVI 778
QY 492 ----- 493
Db 779 SSGDVEKDRVKFHSKQNVLEDAKLAELFKGVRGRGYIEVSGKISHAIPVLAEGKR 838
QY 494 ----- 493
Db 839 IPEFIFTSPMDIKVAPLKLINGNABELTFTSKELLVNLILLNSIGVSDIKIEHKG 898
QY 494 ----- 493
Db 899 YRVYINKSSNGDIVLDSVESIEVEKVEGYVDVSVEDNENFLVGFGLLYAHNSYGY 958
QY 501 GYPKARWYKCECAESVTAWGRHYIEMTKEIEBEKFGKVLVYADTDGFYATIPGKXP-ETI 559
Db 959 GYAKARWYKCECAESVTAWGRQYIDLVRRELEAR-GFKVLYITDGLYATIPGKDWBEV 1017
QY 560 KKKAKEFLKYINSKLPGLLELEYEGFYLRGFFVAKRYAVIDEGRIITRGLGVRRDWS 619
Db 1018 KRRALFEVDYINNSKLPGLLELEYEGFYLRGFFVAKRYAVIDEGRIITRGLGVRRDWS 1077
QY 620 EIAKETQAKVLRAILKEDSVKAVEIVKDVVEIAKYQVPLEKVIHQITKDSLEYKAI 679
Db 1078 EIAKETQAKVLRAILKCHGVNEAVKIVKEVTEKLSYIEPPKLL 1137
QY 680 GPHVAIAKRAAGIKVRPCTIISYIVLRGSGKISDRVILLSEYDPPKHKYDPPYINQ 739
```

Db 1138 GPHVAVAKLMARGIKVKGWIGVILVRGDPISKRAISIEEFPDRKHKYDAEYIENQ 1197
QY 740 VLPVAVRILAEGRKEDLYQSSKQVGLDAWLK 773
Db 1198 VLPVAVRILAEGRKEDLYQSSKQVGLDAWLK 1231
RESULT 7
S71551
DNA-directed DNA polymerase (EC 2.7.7.7) KOD, intein containing precursor - Pyrococcus s
N; Contains: DNA endonuclease (EC 3.1.-.-) PI-I; DNA endonuclease (EC 3.1.-.-) PI-II; DNA
C; Species: Pyrococcus sp.
A; Variety: strain KOD1
C; Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 20-Apr-2000
C; Accession: S71551
R; Kakinaka, H.; Takagi, M.; Imanaka, T.
submitted to the EMBL Data Library, March 1994
A; Description: Cloning and analysis of the DNA polymerase gene from a new hyperthermophi
A; Reference number: S71551
A; Accession: S71551
A; Molecule type: DNA
A; Residues: 1-1670 <KAK>
A; Cross-references: EMBL:D29671
C; Function: <NUCL>
A; Note: as DNA-directed DNA polymerase, catalyzes the polymerization of DNA at th
A; Note: DNA-directed DNA polymerase KOD
C; Function: <EN1>
A; Description: as DNA endonuclease PI-I, catalyzes the hydrolysis of internal phosphodie
A; Note: DNA endonuclease PI-PspI
C; Function: <EN2>
A; Description: as DNA endonuclease PI-II, catalyzes the hydrolysis of internal phosphodi
A; Note: DNA endonuclease PI-PspII
C; Superfamily: DNA-directed DNA polymerase KOD
C; Keywords: endonuclease; hydrolase; nucleotidyltransferase; protein splicing
F; 1-406/767-851/1388-1670/Product: DNA-directed DNA polymerase KOD #status predicted <MA
F; 1-406/Domain: DNA-directed DNA polymerase KOD extein 1 #status predicted <XT1>
F; 407-766/Product: DNA endonuclease PI-I (pol KOD extein 1) #status predicted <MAT2>
F; 767-851/Domain: DNA-directed DNA polymerase KOD extein 2 #status predicted <XT2>
F; 852-1387/Product: DNA endonuclease PI-II (pol KOD extein 2) #status predicted <MAT3>
F; 1388-1670/Domain: DNA-directed DNA polymerase KOD extein 3 #status predicted <XT3>
F; 406-767/Cross-link: peptide (Arg-Ser) #status predicted
F; 851-1388/Cross-link: peptide (Asn-Ser) #status predicted
Query Match 68.7%; Score 2764.5; DB 2; Length 1670;
Best Local Similarity 35.9%; Pred. No. 4.5e-134;
Matches 599; Conservative 83; Mismatches 88; Indels 899; Gaps 4;
QY 1 MIFDTYITKDGKPIIRIFKKEGFEKIEIDPHFPQPIYIALLKDDSAIDBIKAKGERHG 60
Db 1 MILDITDITDGKPIIRIFKKEGFEKIEYDRTPEYFYALLKDDSAIEVVKITAERHG 60
QY 61 KIVRVDAVKKKFLGRDVEVWKLIFEPQDVPAIRGKIREHPAVIDIYEYDIPFAKRY 120
Db 61 TVVTVKRYKVKQKFLGRPVWVKLYFTHPQDVPAIRDKIREHGAVIDIYEYDIPFAKRY 120
QY 121 LIDKGLIPEWGEDEELKMAFQIETFYHGEDEFGKEIMISYADEEARVITWKNIDLPY 180
Db 121 LIDKGLVPEWGEDEELKMAFQIETLYHGEDEFGAEPILMISYADEGARVITWKNVDLPY 180
QY 181 VDVVSNREMIKRFVQIVREKDPDVLITYNGNDPLPYLIGRAEKLGVTLILGRDKHEPE 240
Db 181 VDVVSTEREMIKRFLRVVKEKDPDVLITYNGDNDFAYLKRCCKEKLINFALGSDGS-E 238
QY 241 PKIHRMGDSFAVEIKGRHFDLPVVRRTINLPTVTLAEVAVLGTKSKLGAEEIAAI 300
Db 239 PKIQGRDFAVEVKGRIHFDLPYVIRITINLPTVTLAEVAVFGQPKVAAEITPA 298
QY 301 WETESMKLAQYSMEDARATYELGKEFFPMELAKLIGOSVMDVSRSSSTGNLVEWYLL 360
Db 299 WETGENLVARYSMEDAKVYELGKEFLPMEAQLSRIGOSLMDVSRSSSTGNLVEWFL 358
QY 361 RVAYERNELAPNKDEEERRLRTTYLGGYVKEPERGLWENITVLDPRC----- 410

Db 359 RKAYERNELAPNKDEKELARR-RQSYEGYVKEPERGLWENIVYLDPRCHPADTKVVKV 417
QY 411 ----- 410
Db 418 GKGIINISEVQEGDYVLGIDGMQVRVKWWEYDYKGBELVNINGLCKTPNHKLFPVTKNERQ 477
QY 411 ----- 410
Db 478 TRIEDSLAKSFLTKVKYKGIITTPFVEIGRATSENIPEEVLKGBELAGILLAEGLLRK 537
QY 411 ----- 410
Db 538 DVEYFDSSRKGRIRISHQYRVEITIGKDEEFRDRIITYPERLFGITPFSISEKKGTVAVL 597
QY 411 ----- 410
Db 598 KVAKKNVYLKVEIMDNIESLHAPSVLRGFFEGDGVNRRVSIVATQGTQKNEWKILVS 657
QY 411 ----- 410
Db 658 KLLSQLGIPHOTVITYQYQENGKDRSYILEITGKGLILFQTLIGFISERKNALLKAIS 717
QY 411 -----LYPSIIIVTN 420
Db 718 QRENNLENGFYLSFNFVSTVEYEGKVYDLTLLEGTPYFANGILTHNSLYPSIIITHN 777
QY 421 VSPDTEREGCKNYDVAPIVGYKCPDPPGIFPSIIGELITMQEIKKKAKAIDPIEKX 480
Db 778 VSPDTLAREGCKEYDVAPQVGHRECKDPGPFIPSLGLDLEERQKIKKKAKATIDPIERK 837
QY 481 MLDYRQRAVKLHA----- 493
Db 838 LLDYRQRAKILANSILPEHLPVLERGEVHFVRIGELIDRMMEENAGKVKREGEVLE 897
QY 494 ----- 493
Db 898 VSGLEVPSPNRRTNKALRKVKALIRHDYSGKYVITRLKSGRRKITSGHSLFSVRNGEL 957
QY 494 ----- 493
Db 958 VEVTGDELKPGDLVAVPRRLBELPERNHVLNVELLGTPEETLDIVMTIPVKCKQFFK 1017
QY 494 ----- 493
Db 1018 GMLRTLRFEGEKEPRTRARYLRHLEDLYVRLKIGYEVLDWDSLNKRYRLYEALVEN 1077
QY 494 ----- 493
Db 1078 VRYNGNKEBYLVEFNSIEDAVGIMPLKELKEWKIGTLNGFRMRKLIIEVDESLAKLLGYV 1137
QY 494 ----- 493
Db 1138 SEGYARKQNPKNQWSYVKLYNEPDEVLDDMERLASRFFGKVRGRNRYVEIPKKIGYLL 1197
QY 494 ----- 493
Db 1198 FENMCGVLAENKRIPFVFTSPKGVRLAFLEGYSAMATSTEQETQALNEKRALANQLVL 1257
QY 494 ----- 493
Db 1258 LLNSVGVSAVKLGHDSGVYRVINEELPFVKLDKKKNAYSHVPIKEVLSEVFGVKFQKN 1317
QY 494 ----- 493
Db 1318 VSPQTRKMWEDGRDLPKQAQRLSWLIEGDVVYLDREVSDYDGYVYDLSVENENFL 1377
QY 494 -----NSYGYMGYPKARWYSKECAESVTANGRHYIENTIKEIEBEKFGKLYADT 544
Db 1378 VEGFLVAINSYGYGVARARWYCKEASVTANGREVIITWIKIEIEKYGKVIYSDT 1437
QY 545 DGFYATIPGKEPETIKKAKFLKYNISKLPGILLEFEYEGFVLRGFFVAKRYAVIDSEG 604
Db 1438 DGFATIPGADAETVKKAMEFLNINAKLPGALEFEYEGFYKRGFFVTKKAYVIDSEG 1497

QY 115 PFAKRYLIDKGLIPM-----EGDE-----ELKMAFDIETFYHEGDFGKE-IIMIS 161
 Db 124 PFAKRYLIDNEIIPMTYDFENKQVSIIPKLSVAFDMEVNDRDTEPNPDRDILMAS 183
 QY 162 YADEEARVITTKNIDLPYDVVSNREMIKRFQVIREKDPDVLITTYNGDNFDLPYLK 221
 Db 184 FWDENGKVITYEFNHPNIEVUVNKEKLIKLIETLKEY--DVIITYNGDNFDLPYLKA 241
 QY 222 RAELKGVTLILGRDKEHPEKIHMRMGDSFAVEIKGRIHFDLPVVRTINLPTYLEAVY 281
 Db 242 RAKIYDIDINLKDGE--ELKIRGGMEVRSYIPGRVHIDLVPISRLKLKTYLLEDVY 299
 QY 282 EAVLGKTKSLGABEIAAIWETESMKLAQYSMEDARATYELGKFFFPMEAEALAKIQ 341
 Db 300 YNFGIEKLPHTKIVDYVAND--KTLIEYSLQAKYTKYKGYFFPLEVMSFVNVQ 357
 QY 342 SVMDSVRSSTGNLVEWYLLVAVARNELAPNKDPDEBYRRRLTTLTGGYVKEPPEGLWE 401
 Db 358 TPPEITRMSSQGVYLLMKRAFKNMIVENKDPDEBYRRRLTTLTGGYVKEPPEGLWE 417
 QY 402 NITVLDPRC----- 410
 Db 418 DILSMDFRCHPKGTQVVKGVINIEDVKEGNYVLGIDQKQVKKVKEVEGELINVN 477
 QY 411 ----- 410
 Db 478 GLKCTPNHKIPLRYKIKKXKINKNDYLVROIYAKSLTLTKFGEGKILCKDPETIGNYEK 537
 QY 411 ----- 410
 Db 538 YINDMDEDFILKSELIGILAEGLLRDIEYFDSRGKKRIHQYRVEITVNEDEKDFI 597
 QY 411 ----- 410
 Db 598 EKIKYIFKLPNELYVRRKGTAKITLGCACKDIYKIEILLKQNKYKLPNAILRGFFE 657
 QY 411 ----- 410
 Db 658 GDGYNTVTRAVVNVQNTNNDKIFIASLLDRIGIKYSFYTSYBERGKLRVYEIF 717
 QY 411 ----- 410
 Db 718 SKGDLKPSILISFISRRKNLLNEIIRQTKLYKIGDYGVYDLDDVCVLESYKGEVYDL 777
 QY 411 -----LYPSIIVTHNVSPDTEREGCKNYDA-PIGVKFKCPFGF 451
 Db 778 TLEGPPYFANGILTHNSLYPSIISYNISPTLDCOCK--DVSEKILGHWFCKKEGL 835
 QY 452 IPSILGELITRQEIKKMK--ATIDPI--EKQMLDYRQRAVKLHA----- 493
 Db 836 IPKTLRNLJERRINIKRKKWMAEIGEINEEVNLLDVEQSKILANSILPDEYLTIEE 895
 QY 494 ----- 493
 Db 896 DGIKVVKIGEYIDLLMRKHDKIKFSGISEILETKMLTFSDFDKITKKCEIKKVALIRH 955
 QY 494 ----- 493
 Db 956 PYFGKAYKILRSRTIKVTRGSLFKYENGKIVEKGDDVRFGLIIVPKLTCVDKEV 1015
 QY 494 ----- 493
 Db 1016 VINIPRLINADEEIKDLVITKHDKAFFVKLKTLEDIENKUKVIFDDCILYKELG 1075
 QY 494 ----- 493
 Db 1076 LIDYNTIINKVNDIKILDEEKYAKYFDVIBHGNFKKRCNIQIKIKDYIANIPD 1135
 QY 494 ----- 493
 Db 1136 KEPEDCEIGAYSGKINALLKDEKLAKFLGFFVTRGRLLKQKLKGETVYISVYKSLPEY 1195

QY 494 ----- 493
 Db 1196 QKEIAETFEVFGAGSVKDKVTMDNKIVVLVLYIFKCGDKDKKHIPBELFLASBSVIK 1255
 QY 494 ----- 493
 Db 1256 SFLDGLKAKNSHKGTSTFMAKDEKYLNMILFNLVGIPTRTPVONKGYKLTLPKY 1315
 QY 494 -----NSYGYMGYPKARWYSKE 511
 Db 1316 GTVXDLMDLVEKETEAFESYGVYVYDLSVEDNENFLVNNIVAHNSVGYLAFPRARFYSRE 1375
 QY 512 CAESVTAWGRHYIMTKEIEBKFGFKVLYADTDGFIATIPGEK--PETIKKAKPELKY 569
 Db 1376 CABITVTLGRKYILLETYKE-AEKFGFKVLYDIDGFFYA--IWKKISKEELIKKAMEFVY 1433
 QY 570 INSKLPGLLELEYEGFYLRGFFVAKRYAYVIDEGRITTRGLEVVRDSEIATKETOAKY 629
 Db 1434 INSKLPGLTMELEFGYFKRGIFVTKRYALIDENGRVTVKGLFVRDWSNIAKITQRRV 1493
 QY 630 LEALIKEDSVKAVEIKVDVVERIAKQVPLEKLVIEHQTKDLSYKATGPHVAIAKRL 689
 Db 1494 LEALLIVEGSIKAKKIIQDVIKDLREKKIKEDLIYQTJTKDPKEYTTPHVEIAKRL 1553
 QY 690 AAKGKIVRPGTIIISVILRGSKISDRVILLSEYDPKHKHYDDPYIENQVLPVLRILE 749
 Db 1554 MREGRIKVGDIIGYIIVKGTKSISERAKLPEEVD--IDDIDVNYIDNQILPPVLRIME 1611
 QY 750 AFGYKEDKLYOSSKQVGLDAWLK 773
 Db 1612 AVGVSKNELKEGA-QLTLDKFFK 1634

RESULT 10
 C69028
 DNA-dependent DNA polymerase family B (PolB1) - Methanobacterium thermoautotrophicum (C) Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: C69028
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, I.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 175, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fun
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: C69028
 A:Status: praliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-586 <MTH>
 A:Cross-references: GB:AE000888; GB:AE000666; NID:G2622304; PIDN:AA85697.1; PID:G2622
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1208

Query Match 30.0%; Score 1209.5; DB 2; Length 586;
 Best Local Similarity 42.9%; Pred. No. 7.7e-55;
 Matches 255; Conservative 107; Mismatches 176; Indels 57; Gaps 10;

QY 1 MIFDPTDITKDGKPIIRIFKE---NGEFKIELDPHQPIYIALLKD-DSAIDEIKAKG 56
 Db 7 VLLDIDVTVYVDEVPVIRLFQKDKSGNEPIIAHDSFRFYIATPTDLDCELELEEL 66
 QY 57 ERHGKIVRVVDVAVKVK-KFLGRDVBVWKLITEHPQDVPALRGKIRHPAVIDIYEYDIP 115
 Db 67 EK-----LEVEMEDLGRPTVIRIEFRHPQDVPKIRDIRDLESVRDIREHDIP 116
 QY 116 FAKRYLIDKGLIPME-----GDEELKLMARFI 142
 Db 117 FYRRLIDKISIVMEELFQGVSDSAPSTVTDRTVETVGRVQSTGSAHGLDILSFDI 176
 QY 143 ETPYHEG-DFGKGEIIMISYA--DEBEARVITWKNDLPVDVWSNERMIKRFQVIRE 200
 Db 177 EVRNPHGMPPDEKDEIVMIGVAGNMGVESVISTAGDHLDFVEVVEDERELLERFAFVID 236

Db	226	DRVLREFVYRAAFDPDIIVGYNHDFWPLYMERARRLGKLDVTR-RVGAETPTSV	283
Qy	246	MGDSFAVEIKGRIHFDLFPVVRITNLPVTLEAVEAVLGTCKSLGAE--IAAIWE	302
Db	284	YGH---VSVQRLNVDVDAEMPELIMKTLEVAASYLGWVKSERVIEWRIPEYWD	340
Qy	303	TEESMKLAQYSMEDARATYELGKEFPFMEAEAKLIGOSVWDVSRSTGNLVEWYLLRV	362
Db	341	DEKKQLLERYALDDVRATYGLAEKMLPFAQLSTVTGVPDQVGAMGVGFRLEWYLMRA	400
Qy	363	AYERNELAPNKPDEEYRRRLRTTYLGGYKKEPERGLWENITVLDPRCLYPSIIVTHNVS	422
Db	401	AYDNMELPNRVE-----RGESEYKGAIVLPKUGVHENVVLDFFSMTPSIIMIKYNG	454
Qy	423	PDTL--ERECKNYD---VAPIVGYKFCDDPGFIPSIILGELITMRQEIKKKKA-TIDP	476
Db	455	PDTIVDDSECPKYGGCVVAPEVGHRRPRSPGPFKTVLENLLKLRQVKEKKEKFPDPS	514
Qy	477	IEKMDLRQRAVKLHANSYGYMGYPKARWYKCAESVTAMGRHVIENTIKEIEBKFG	536
Db	515	PEYRLYDERQALKVLNANASYGGMGSHARWYCKRCABAVTAWGNLI-LTAIEYARKLG	573
Qy	537	FKVLVYADTDFVATIPGKPEIKKAKAEFLKYINSKLPGLLELEYEGFYLRGFFV-AKK	595
Db	574	LKVIYDGTDSLFFVVD-----KEKVEKLIEFVEKELG--FEIKIDKIYKKVFFTEAKK	624
Qy	596	RYAVTDEGRITRGLVVRDWSIAKETOAKVLEAILKEDSVKAVEIVKDVVEIAK	655
Db	625	RYVGLLEDGRIDIVGFAVRGDMCELAKEVQEAABIIVLNTGNVDKALSIREVVKOLRE	684
Qy	656	YQVPLEKIVHBOITKDLSEYKAIGPHVAIAKRLAAGIKVRPGTIISYIVLRGSGKISD	715
Db	685	GRVPTIKLIWTKLSRIEYEHADAPHVVAARMEKEAGYEVSPGDKGVYIVKSGSVSS	744
Qy	716	RVLLSEVDPKXKHVDPPYIENQVLPVLRILEAFGRKEDLKYSQSKQVGL	768
Db	745	RAYPFVMDPS--TIDVNYIDHQIVPAALRIILSYFGVTREKQLKAAATVQVRS	795
RESULT 13			
E72515			
probable DNA-directed DNA polymerase APE2098 - Aeropyrum pernix (strain K1)			
C:Species: Aeropyrum pernix			
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000			
C:Accession: E72515			
R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah			
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K			
DNA Res. 6, 83-101, 1999			
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr			
A:Reference number: A72450; PMID:99310339; PMID:10382966			
A:Accession: E72515			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-784 <XAW>			
A:Cross-references: DDBJ:AP000063; NID:g5105654; PIDN:BAA81109.1; PID:g5105797			
A:Experimental source: strain K1			
C:Genetics:			
C:Superfamily: herpesvirus DNA-directed DNA polymerase			
Query Match 28.3%; Score 1138; DB 2; Length 784;			
Best Local Similarity 34.6%; Pred. No. 5.3e-51;			
Matches 282; Conservative 137; Mismatches 287; Indels 110; Gaps 19;			
Qy	11	DGKPIIRIFKNGEKFELDPHFQYIYVALLKDDSAIDEIKAKGERHGKIVRVVDAVK	70
Db	16	DGSRVVFY----GEFR-----PFYFVLPDGSVGLDQLAAMIRRLSRPSSPILVER	63
Qy	71	VKKKELGRDVEVWKLIFEPHQDVPALRGKIREHPAVIDIVEYDIPPAKRYLIDKGLIPM-	129
Db	64	VRRRFIGREVALKVTILVPASVREYREAVRRLGGVDRVLEADIPPALRFIIDFNLYPMR	123

Qy	130	-----EGD-----BELKLMADITFFYH-EGDEF	152
Db	124	MYVAEVRVAVPHGYSDRAYTLTSGDITREDETRIQEDPLKGLRVNAPDIEVYSKORTDDP	183
Qy	153	KGGEIIMISYAD-----EEEARVITWKNIDLPYVDVWSNEREMIKRFYQIVREKDPD	204
Db	184	KDPVIMIGLQOAGGIEIILEAEDR-----SDKKVIAGFVERVKSIDPD	227
Qy	205	VLITVNGDNFDLPYILKRAEKLGVTLGRDKEHPEPKHRMGDSFAVEIKGRIHFDLFP	264
Db	228	VTVGYNQNRFDWPFYUVERARVLGVKLVAGRRSVPEQPLYG-----HYSVSGRLNVDLDD	282
Qy	265	VVRRTINLPVTLEAVEAVLGTCK--SKLGAE--EIAAIWETEBESMKKLAQYSMEDARA	320
Db	283	FAEELHEVVKVTLVEAD-YLGVYKIGERVLTWQICEYWDPSKRLAKYLRDDVRS	341
Qy	321	TYELGKEFPFMEAEAKLIGOSVWDVSRSTGNLVEWYLLRVAYERNELAPNKPDEEYR	380
Db	342	TMGLAEKFLPFCBELSQVSLFDQVMAASVGFLEWRLIREAAKGLSLVNRVERSEGR	401
Qy	381	RLRLTYLGGYKKEPERGLWENITVLDPRCLYPSIIVTHNVSPTDLREGECKNYD-----	435
Db	402	-----YAGAIVRFPKPGVHEDIAVLDFAVMYFNIMVKYNGVDPDLVRFP-EEYGEVEY	454
Qy	436	VAPIVGYKFCDDPGFIPSIILGELITMRQEIKKKKA-TIDPIEKMDLYQRAVKLHAN	494
Db	455	TAPEVGHKPRKPPGPFKKILERFLSRQIRSENMKCHPPDPSPEYKLLDERQKAIKLAN	514
Qy	495	SVYGYMGYPKARWYKCAESVTAMGRHVIENTIKEIEBKFGKVLXADTDFGYATIPGE	554
Db	515	ASVGYMGWPHARWYCRECAEAVTAMGRSITIRAIKAGE-LGLEVIYDGTDSLFLVKNDEP	573
Qy	555	KPETIKKAKFEIKYINSKLPGLLELEYEGFYLRGFFV-AKRYAVIDEEGRIITRGLEV	613
Db	574	KVERLIRFVEELGF-----DIKVDKVVRRVFFTEAKRYVGLTVDGRKIDVVGFEA	624
Qy	614	VRDWSIAKETOAKVLEAILKEDSVKAVEIVKDVVEIAKYOVPLEKVIHQITKDL	673
Db	625	VRDWSELAKETQFVRAEIVLKTGSDVEADVVRNIIKRLRGQVDMKRLVWKILTRPP	684
Qy	674	SEYKAIGPHVAIAKRLAAGIKVRPGTIISYIVLRGSGKISRIVILLSYDPKHKYKDPD	733
Db	685	SMYEAQPHVTAALLMERAGIKVEPGAKIGYVTVKSGPLVTRA--KPYFMASKEEVDVE	742
Qy	734	YYTENQVLPVLRILEAFGRKEDLKYSQSKQVGLD	769
Db	743	YYVDKQVPAALRILOQYFVGTREKRLKGGGRQSTLLD	778
RESULT 14			
T05731			
DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - soybean			
C:Species: Glycine max (soybean)			
C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 18-Jun-1999			
C:Accession: T05731			
R:Collins, J.T.B.; Cannon, G.C.; Heinrichst, S.			
submitted to the EMBL Data Library, August 1997			
A:Reference number: Z15439			
A:Accession: T05731			
A:Status: preliminary; translated from GB/EMBL/DDBJ			
A:Molecule type: mRNA			
A:Residues: 1-1088 <COL>			
A:Cross-references: EMBL:AF020193; NID:G2895197; PIDN:AAC18443.1; PID:g2895198			
C:Genetics:			
C:Gene: Pol delta			
C:Function:			
A:Description: catalyzes replication of DNA			
C:Superfamily: herpesvirus DNA-directed DNA polymerase			
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase			
Query Match 17.2%; Score 693.5; DB 2; Length 1088;			
Best Local Similarity 27.2%; Pred. No. 5.5e-28;			
Matches 237; Conservative 149; Mismatches 324; Indels 161; Gaps 30;			


```
QY 15 IIRIFKKGEPKIEIDPH-FQPIYIALLKODSAIDEI-----KAIGERHGK 61
Db 88 IIRIFGVTKGHSVCNCHGPEFYICCPGGMGDDISHFQHTLEGEMREANNSVGK 147
QY 62 IYRVVDANKV-----KKFLGRDVEVWKLIFEPHQDPVALRGKIRHPDAVID--- 108
Db 148 FVRLEWQRRSIMYQOSNSOPFL-----KIVVALPTWVASCGR-ILDRGQLDGLG 199
QY 109 -----IYEDIPFAKRYLIDKGLI----- 127
Db 200 MKSFLTYESNVLFALEFMDCNIVGNWIGIPAGKYKTKAKSLSCVQLEFDCLSSELISH 259
QY 128 PNEG-----BELKMAFDIEFYHGG--DEFGKBEIIMISYADEBEARVITWKNIDIPY- 180
Db 260 APEGYSXWAPRILISFDICAGRGHPPEPTHDPVQI-----ANLVTLOGEDQFFI 312
QY 181 -----VDVV--SNERMIKRFVQIVREKDPDVLITVNGDNFDPVLYIKRAE 224
Db 313 RNWMTLKSQSPVGVDMVPFETEREVLLAWRDFREVDPDIIIGNYICKFPLDPLIALERAL 372
QY 225 KLGVT--LLGRDKHP-----EPKIHMGOSFAVEIKGRHFDLPFVVRTINLPT 274
Db 373 NLKIAEFILGRINSRVKVDKTTFSRQYGTRESKEVAVEGRVTFDQLQWQDYKLSS 432
QY 275 YLEAVYEAVLCKTKSLGABEIAAWE--TESMKKLAQYMEDARATYELGKE--FPFM 331
Db 433 YLSNVSSHFLSEQEDVHHSIISDLQNGNAETRRRLAVYCKDAYLPORLLDKMFIYN 492
QY 332 EABLAKLIGQSV-MDVSSSTGNVWYLLRVAVYERNE LAPN-KPDDEEYRRRLRTTYLG 389
Db 493 YVEARVTVGVPISFLSRGQSIKVL--QLLRARQKNLVIENAKOAGSE-----QGTFFG 546
QY 390 GVYKEPERGLWEN-ITYLDFRCLVPSIIVTHNVSPDTLE-REGKNYDVAP-----IVG 441
Db 547 ATVLEARAGFYEKPIATLDFAFLPSINWALNVCVLVIPEDARKNIPESVNRTPSG 606
QY 442 YKPKCKA-PFQFIPISILGELITMRQEIKKQKATIDPIEKKMLDYORAVKLANHANSYGY 499
Db 607 ETVKSNLQKGLPELLELTARKAKADLEAKDPLEKAVLDGRQLALKISANSYGF 666
QY 500 MGYPKARWYSECAESVTAGSHVIEMTIKIEKF-----CF-----KVLADTDGFYATI 551
Db 667 TGAIGQPCLEIISSVTSYGRQWIEHTKLVEDKFTLNGYERNAEVIYDGTDSVMVQF 726
QY 552 PGEKPEIKKAKFLKINSKLPGLLELEYEGFLRGFFVAKRYAVI-----DESGRI 606
Db 727 GVSAAVEAMNLGREAHEISGTFKPKLEFEKYYPVLLISKRYAGLFWTKPDNFDKM 786
QY 607 TTRGLEVVRDWESEIAKETAQVLEALKEDSVSEKAVEIKVDVVEEIAKYQVPLEKLVH 666
Db 787 DTGKLETVRDNCLLVKNLNDCLHKILIDRIDPCAVQYVKNASIDLLMNRMDLSLVIT 846
QY 667 EKITKDLSEYKAIGHVAIAKRLAAKGIKVRP--GTTISYIVLRGS-----GKISDRVI 718
Db 847 KGLTKTGDDYEVKAAHVELAERMRKDAATAPNVGDRVPYVIIKAAKGAKAYESEDPIY 906
QY 719 LLESDPKKHYKDPDYIENQVLPFAVRILE 749
Db 907 VLENNP-----IDPHYYLENQISPKILRIFE 933
```

```
RESULT 15
S75407
Probable DNA-directed DNA polymerase (EC 2.7.7.7) - Sulfolobus solfataricus
N:Alternate names: protein c04041
C:Species: Sulfolobus solfataricus
C>Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C:Accession: S75407
R:Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A:Title: Organizational characteristics and information content of an archaeal genome: 1
A:Reference number: S73076; MUID:97055432; PMID:8899719
```

```
A:Accession: S75407
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-764 <SEN>
A:Cross-references: EMBL:Y08257; NID:g1707772; PID:g1707813
A:Experimental source: strain P2
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: nucleotidyltransferase
```

```
Query Match 16.8%; Score 676.5; DB 2; Length 764;
Best Local Similarity 25.3%; Pred. No. 2.6e-27;
Matches 206; Conservative 168; Mismatches 312; Indels 127; Gaps 24;
```

```
QY 2 IFDTYITKDGKPIIRIFK-KENGEPKIELDPHFQPIYIALLK--DDSAIDKAIKGR 58
Db 7 ILDFSVEIKGNPLVYIWSVDDEGNSVVIDNNFPFYIIVGEGNEIENIK----- 60
QY 59 HGKIVRVDAVKKFLGRDVEVWKLIFEPHQDPVALRGKIRHPDAVIDIYEYDI--- 114
Db 61 --KNCEALQITVKRYKYLGNIVDALLIQSTPTQIKCKREKISELNNIKGIFDADIRYTM 118
QY 115 -----PFA-----KRYLDKGLIPMEGD--BELKMAFDIETFFH 147
Db 119 RYSLDFCLPPTWFAFVNEVDFGRTKKAIVLDKILSHYEGNMPFLRTIGVDFQIY-- 176
QY 148 EGDEFG-----KBEIMISYADEBEARVITWKNIDLPYVDVVSNEREMIKRFVQIVREK 202
Db 177 --SKYSLNPRKDPVIMSLWSKEGMPQS-----LDEGIDDLKIIRRFVDYILNYD 226
QY 203 PDVLITNGDNFDPVLYIKRAEKLGVTLGLGRDKEHPKIRHMGDSFV-----EIKG 256
Db 227 PLIIFVYSDLLPWKVIITERRASSLGKVIDIGR-----KIGSEVSVGYGHSYISG 276
QY 257 RIHFDLPVVRRTINLPTVYLEAVYEAVLGKTKSLGAB--EIAAIWTEESMKKLAQYS 314
Db 277 RLNVDLTGLLVNERSLSGHVDLIDV--SNYLGISPSYSFKWYEISRYWDNKNRIIREYS 335
QY 315 MEDARATYELGKEFPFMEAEIAKLIGQSVWDVSRSTGNLWYLLRVAVYERNE LAPNKP 374
Db 336 IENASIIYLLGNILYSTSELVKIVGLPLDKLSVASWGNRIETSLIRTATSKSGELIPIRM 395
QY 375 DEEYRRRLRTTYLGYKVEPERGLWENITYLDFRCLVPSIIVTHNVSPDTLERSGCKNY 434
Db 396 DNPNRPSIKK--NIIQPKVGITYDVYVLDISSVYSLVRKENIAPDLTVKQCCDC 451
QY 435 DVAPIVGKFCDFPGFIPSIILGELITMRQEIKKQKATIDPIEKKMLDYORAVKLAN 494
Db 452 YSSPISNRYKREPSGLYKTFDLSNVRDSNKIV-----IEELI-----S 493
QY 495 SYGYMGYPKARWYSECAESVTAGRHVIEMTIKIEIEKFGFKVLYADTDGFIATIPGE 554
Db 494 SFNDYVHWVNRWYREIATASAFDEFSNELIRFIIDLIKSS--GLDIVLANDLLIFV--- 548
QY 555 KPETIKKAKPELKTINSKLPGLLELEYEGFLRGFFVAKRYAVIDEGEITTRGLEV 614
Db 549 --GGSRDKNVLIITKINLYN--LDVKVKIFYKSLVLDNNRYAGLSGDKI-----DIA 599
QY 615 RR-----DWSEIAKETQAKVLEALKEDSVSEKAVEIKVDVVEEIAKYQVPLEKLVH 669
Db 600 RKEEDMNLCELARNIKKIIEELISKDVKAKIKLVASTWIKLRGDFNEELITWAKI 659
QY 670 TKDLSEYKAIGHVAIAKRLAAKGIKVRPGTIIISYIVLRGSKTISDRV---ILLSEYDPK 726
Db 660 ERDLNENYQLPFVTAARKAIQSGYLISKDSKIGYIVKVLGGLPLNDRAEPFFLVE--- 715
QY 727 KHYKDPDYIENQVLPFAVRILEAFGYRKEDLK 759
Db 716 KNRIDIEYVD-QIFRETILKLLKPLGVNNEESLK 747
```

```
Search completed: February 9, 2004, 11:53:19
Job time : 61 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 11:39:55 ; Search time 23 Seconds

(without alignments)
1582.551 Million cell updates/sec

Title: US-09-803-165-34

Perfect score: 4026

Sequence: 1 MFEDTYTKDGKPIIRIFK.....KEDLYQSSKQVGLDAWLKK 774

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3446.5	85.6	1829	1 DPOL_THEST	O33845 thermococcu
2	3356	83.4	1702	1 DPOL_THELI	P30317 thermococcu
3	3253.5	80.8	773	1 DPOL_THES9	P56689 thermococcu
4	3204.5	79.6	775	1 DPOL_THES9	Q56366 thermococcu
5	3164.5	78.6	771	1 DPOL_PYRAB	P77916 pyrococcus
6	3151	78.3	775	1 DPOL_PYRFU	P80061 pyrococcus
7	2934.5	72.9	1312	1 DPOL_PYBSD	Q51334 pyrococcus
8	2886	71.7	1235	1 DPOL_PYRHO	Q59610 pyrococcus
9	2809	69.8	1523	1 DPOL_PYRHO	P74918 thermococcu
10	2782	69.1	1671	1 DPOL_PYRHO	P77933 pyrococcus
11	2769.5	68.8	1699	1 DPOL_THES8	Q9hh84 thermococcu
12	2535.5	63.0	1668	1 DPOL_THES9	Q9hh05 thermococcu
13	1392	34.6	824	1 DPOL_METVO	P52025 methanococc
14	1301	32.3	781	1 DPOL_ARCFU	Q29753 archaeoglob
15	1292	32.1	1634	1 DPOL_METUA	Q58235 methanococc
16	1209.5	30.0	586	1 DPOL_METUA	Q27276 methanobact
17	1138	28.3	784	1 DPOL_AERPE	Q93746 aeropyrum p
18	697.5	17.3	1105	1 DPOL_ORYSA	Q91r66 oryza sativ
19	693.5	17.2	1088	1 DPOL_SOYBN	O48901 glycine max
20	691	17.2	763	1 DPOL_SULSH	O05706 sulfolobus
21	676.5	16.8	764	1 DPOL_SULSO	P95979 sulfolobus
22	676.5	16.8	1081	1 DPOL_ARATH	Q91v77 arabidopsis
23	657.5	16.3	1038	1 DPOL_CANAL	P46588 candida alb
24	657.5	16.3	1086	1 DPOL_SCHPO	P30316 schizosacch
25	624.5	15.5	1092	1 DPOL_DROME	P54358 drosophila
26	621.5	15.4	1097	1 DPOL_YEAST	P15436 saccharomyc
27	619.5	15.4	1103	1 DPOL_RAT	O54747 rattus norv
28	616.5	15.3	872	1 DPOL_SULOH	O50607 sulfurispha
29	616.5	15.3	882	1 DPOL_SULSO	P26811 sulfolobus
30	616.5	15.3	1105	1 DPOL_MOUSE	P52431 mus musculus
31	615.5	15.3	1103	1 DPOL_MESAU	P72823 mesocricetu
32	614.5	15.3	1094	1 DPOL_PLAFK	P20315 plasmodium
33	611	15.2	1107	1 DPOL_HUMAN	P28340 homo sapien

34	609	15.1	1513	1 DPOA_OXYTR	Q27152 oxytricha t
35	605.5	15.0	1106	1 DPOL_BOVIN	P28339 bos taurus
36	587	14.6	1492	1 DPOA_OXYNO	Q94636 oxytricha n
37	584	14.5	1451	1 DPOA_RAT	O89042 rattus norv
38	583.5	14.5	1462	1 DPOL_HUMAN	P09884 homo sapien
39	581	14.4	1081	1 DPOL_CAEEL	P08629 caenorhabdi
40	578.5	14.4	875	1 DPOL_SULAC	P95690 sulfolobus
41	570.5	14.2	959	1 DPOL_AERPE	O93745 aeropyrum p
42	570	14.2	1465	1 DPOA_MOUSE	P33609 mus musculu
43	569.5	14.1	1243	1 DPOA_ORYSA	O48653 oryza sativ
44	548.5	13.6	1339	1 DPOA_TRYBB	P27727 trypanosoma
45	527.5	13.1	1468	1 DPOA_YEAST	P13382 saccharomyc

ALIGNMENTS

RESULT 1

ID	DPOL_THEST	STANDARD;	PRT; 1829 AA.
DT	15-DEC-1998 (Rel. 37, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	DNA polymerase (EC 2.7.7.7).		
GN	POL.		
OS	Thermococcus sp. (strain TY).		
OC	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;		
OC	Thermococcus.		
OX	NCBI_TaxID=110163;		
RN	[1]_SEQUENCE FROM N.A.		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=98094267; PubMed=9434178;		
RA	Niehaus F., Frey B., Antranikian G.;		
RT	"Cloning and characterisation of a thermostable alpha-DNA polymerase		
RT	from the hyperthermophilic archaeon Thermococcus sp. TY."		
RL	Gene 204:153-158(1997).		
CC	!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate		
CC	+ [DNA](N).		
CC	!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES		
CC	A POST-TRANSLATIONAL EXCISION OF THE THREE INTERVENING REGION		
CC	(INTERNS) FOLLOWED BY PEPTIDE LIGATION.		
CC	!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL; Y13030; CAA73475.1; ..		
DR	HSP; P56689; ITGO		
DR	InterPro; IPR006172; DNA_pol_B.		
DR	InterPro; IPR006134; DNA_pol_B_dom.		
DR	InterPro; IPR006133; DNA_pol_B_exo.		
DR	InterPro; IPR003586; Hedgehog_HintC.		
DR	InterPro; IPR003587; Hedgehog_hintN.		
DR	InterPro; IPR006141; Intein.		
DR	InterPro; IPR006142; INTEIN.		
DR	InterPro; IPR004042; Intein_endonuc.		
DR	InterPro; IPR004578; Pol2.		
DR	Pfam; PF00136; DNA_pol_B_4.		
DR	Pfam; PF03104; DNA_pol_B_exo; 1.		
DR	PRINTS; P00379; INTEIN.		
DR	SMART; SM00305; HintC; 3.		
DR	SMART; SM00306; HintN; 3.		
DR	SMART; SM00486; POLSc; 1.		
DR	TIGRFAMs; TIGR01443; Intein_Cterm; 3.		
DR	TIGRFAMs; TIGR01445; Intein_Nterm; 3.		
DR	TIGRFAMs; TIGR00592; pol2; 2.		
DR	PROSITE; PS00116; DNA_POLYMERASE_B; FALSE_NEG.		

DR PROSITE; PS50818; INTEIN_C_TER; 3.
 DR PROSITE; PS50819; INTEIN_ENDONUCLEASE; 2.
 DR PROSITE; PS50817; INTEIN_N_TER; 3.
 KW Transferase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
 KW Protein splicing.
 FT CHAIN 1 409 DNA POLYMERASE, 1ST PART (POTENTIAL).
 FT CHAIN 410 769 INTEIN I.
 FT CHAIN 770 855 DNA POLYMERASE, 2ND PART (POTENTIAL).
 FT CHAIN 856 1392 INTEIN II.
 FT CHAIN 1393 1441 DNA POLYMERASE, 3RD PART (POTENTIAL).
 FT CHAIN 1442 1598 INTEIN III.
 FT CHAIN 1599 1829 DNA POLYMERASE, 4TH PART (POTENTIAL).
 SQ SEQUENCE 1829 AA; 211875 NM; A113A8BC57EB9CB3 CRC64;

Query Match 85.6%; Score 3446.5; DB 1; Length 1829;
 Best Local Similarity 42.2%; Pred. No. 1.2e-179;
 Matches 771; Conservative 0; Mismatches 3; Indels 1055; Gaps 3;

QY 1 MIFPDYITKDGKPIIRIFKXENGEPKIELDPHPQPIYIALLKDDSAIDIKAIKGERHG 60
 DB 1 MILDYITKDGKPIIRIFKXENGEPKIELDPHPQPIYIALLKDDSAIDIKAIKGERHG 60

QY 61 KIVRVDAVKYKKFLGRDVEWKLIFHPQDVPAALRGKIREHPAVIDIYEYDIPPAKRY 120
 DB 61 KIVRVDAVKYKKFLGRDVEWKLIFHPQDVPAALRGKIREHPAVIDIYEYDIPPAKRY 120

QY 121 LIDKGLIPMEGDEELKMAFDIETFYHGEDFGKEIIMISYADEEARVITWKNIDLPI 180
 DB 121 LIDKGLIPMEGDEELKMAFDIETFYHGEDFGKEIIMISYADEEARVITWKNIDLPI 180

QY 181 VDVVSNREMIKRFVQIVREKDPDLITVNGDNFCLPYLIKRAEKIGVTLLGLGRDKEHPE 240
 DB 181 VDVVSNREMIKRFVQIVREKDPDLITVNGDNFCLPYLIKRAEKIGVTLLGLGRDKEHPE 240

QY 241 PKIHRMGDSFAVEIKGRIFHDLFPVVRTINLPTVYTLAEVAVLGTKSKLGAEEIAAI 300
 DB 241 PKIHRMGDSFAVEIKGRIFHDLFPVVRTINLPTVYTLAEVAVLGTKSKLGAEEIAAI 300

QY 301 WETESMKLAQYSMEDARATYELGKEFFPMEAEKLIIGOSVMDVSRSGTGNLVEWYLL 360
 DB 301 WETESMKLAQYSMEDARATYELGKEFFPMEAEKLIIGOSVMDVSRSGTGNLVEWYLL 360

QY 361 RVAYERNEAPNKPDEEYRRRLRTTYLGGYVKEPERGLWENITVLDPRC 410
 DB 361 RVAYERNEAPNKPDEEYRRRLRTTYLGGYVKEPERGLWENITVLDPRC 410

QY 411 411 410

QY 421 GKGIVNISDVKEGDYILGIDGWQVRKVKYHYEGKLININGLKCTPNHKVPVVTENDRQ 480
 DB 421 GKGIVNISDVKEGDYILGIDGWQVRKVKYHYEGKLININGLKCTPNHKVPVVTENDRQ 480

QY 411 411 410

QY 481 TRIRDSLAKSFLSGKVKGKIIITTKLFEKIAEPKKNPSEEBILKGELSGIILAEGTLLRK 540
 DB 481 TRIRDSLAKSFLSGKVKGKIIITTKLFEKIAEPKKNPSEEBILKGELSGIILAEGTLLRK 540

QY 411 411 410

QY 541 DIEYFDSRGRKKRISHQYRVEITIGENEKELLERYIFDKLFGIRPSYKXKGDNTALKI 600
 DB 541 DIEYFDSRGRKKRISHQYRVEITIGENEKELLERYIFDKLFGIRPSYKXKGDNTALKI 600

QY 411 411 410

QY 601 TTAKKAVYLOIBELLKNIESLYAPAVLGRGFFERDATVKNIRSTIVVTOGNNKWKIDIVA 660
 DB 601 TTAKKAVYLOIBELLKNIESLYAPAVLGRGFFERDATVKNIRSTIVVTOGNNKWKIDIVA 660

QY 411 411 410

QY 661 KLKDSLGIPYSRYEYKYIENGKELTKHILEITGRDGLILFQTLVGFISSEKNEALEKAIE 720
 DB 661 KLKDSLGIPYSRYEYKYIENGKELTKHILEITGRDGLILFQTLVGFISSEKNEALEKAIE 720

QY 411 411 410

QY 721 VREWNRLKNNSFNLSTFEVSSEYKGEVYDTLGNPNYPFANGILLTHNSLYPSIIVTN 780
 DB 721 VREWNRLKNNSFNLSTFEVSSEYKGEVYDTLGNPNYPFANGILLTHNSLYPSIIVTN 780

QY 421 VSPDTLEREGCKNYDVAPIVGKFCDFPGFIPSGILGELITMRQEIKKMKATIDPIEKX 480
 DB 421 VSPDTLEREGCKNYDVAPIVGKFCDFPGFIPSGILGELITMRQEIKKMKATIDPIEKX 480

RESULT 2
 DPOL_THELI

DB 781 VSPDTLEREGCKNYDVAPIVGKFCDFPGFIPSGILGELITMRQEIKKMKATIDPIEKX 840
 QY 481 MLDYRQRAVKLHA 493
 DB 841 MLDYRQRAVKLHA 900
 QY 494 493
 DB 901 VDNITAFSLNKESKSEIKVKALIRHKYGEAYEVELNSGRKIITRGHSLFTIRNGKI 960
 QY 494 493
 DB 961 KEIMGEEVKVGDLIIIVPKVKLNEKEAVINIPELISKLPDEDTADVMTTPVKGRKNPFFK 1020
 QY 494 493
 DB 1021 GMLRTLKWI FGEESKRIRTFNRYLFLHEELGFVKLLPRGYEVTDWEGLYRQLYEKLVK 1080
 QY 494 493
 DB 1081 NLRYNGNKREYLVRFNIDIKDSVCPFRKELBEWKIYTKGFRXXCILKVDSDPGFLGY 1140
 QY 494 493
 DB 1141 VSEGYAGAQNKTGCMYSYVKLYNENPNVLKDMKNIAEKFPGKVRVGNKCVDI PKQWYL 1200
 QY 494 493
 DB 1201 LAKSLCGVTAENKRIPSIIFDSSEPVRAWFLRAYFVGDDIHPSKRLSLSTKSELLANQL 1260
 QY 494 493
 DB 1261 VFLLNSLGVSSIKIGFDSGVYRVVYINEDLPFLQTSRQNTVYPNLI PKVELEEIPGRAFK 1320
 QY 494 493
 DB 1321 KNITPEKFKELADSKLDRKVKLLDFLLNGDIVLDRVKNVEKRYEGVYDLSVEDNEN 1380
 QY 494 493
 DB 1381 FLVGFGLLAHNSYGYGMYPKARWYKCEASVTAWGRHYIEMTIKIEEKGFGKLYA 1440
 QY 543 D 543
 DB 1441 DSVTGDTEIIVKRNRIEFVPIEKL FERVYRIGEKEYCILEDBEALTLDRGKLIWKV 1500
 QY 544 543
 DB 1501 PYWRHRAKKVRIWITNSWIDYTDHSLIVAEDGLKEARPMIEGKSLIATKDDLSG 1560
 QY 544 543
 DB 1561 VEYIKPHAIEISYNGYVYDIEVEGTHRF FANGILVHNTDGYATIPGEKPTIKKAKE 1620
 QY 566 FLKYVNSKLPGLLELEVEGFLRGFFVAKRYAVIDEGRITTRGLEVVVRDSEI AKET 625
 DB 1621 FLKYVNSKLPGLLELEVEGFLRGFFVAKRYAVIDEGRITTRGLEVVVRDSEI AKET 1680
 QY 626 QAKVLEAILKEDSVEKAVEIKDVWVEEIAKYQVPEKLVHIEQITKDLSEYKAIGHVAI 685
 DB 1681 QAKVLEAILKEDSVEKAVEIKDVWVEEIAKYQVPEKLVHIEQITKDLSEYKAIGHVAI 1740
 QY 686 AKELAAGIKVRPGTTIISYIVLRGSKISDRVILLSEYDPKXHKYDPDYYIENQVLPAYL 745
 DB 1741 AKELAAGIKVRPGTTIISYIVLRGSKISDRVILLSEYDPKXHKYDPDYYIENQVLPAYL 1800
 QY 746 RILEAFGYRKEDLYOSSKQVGLDAWLKX 774
 DB 1801 RILEAFGYRKEDLYOSSKQVGLDAWLKX 1829

ID DPOL_THELI STANDARD; PRT; 1702 AA.
AC P30317;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) (Vint DNA polymerase) [Contains:
DE Endonuclease PI-Tlii (EC 3.1.-.-) (Tli pol-1 intein) (IVPS2);
DE Endonuclease PI-Tlii (EC 3.1.-.-) (Tli pol-2 intein) (IVPS1)].
GN POL.
OS Thermococcus litoralis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=2265;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92302285; PubMed=1608969;
RA Perler F.B., Comb D.G., Jack W.E., Moran L.S., Qiang B.,
RA Kucera R.B., Benner J., Slatko B.E., Nwankwo D.O., Hempstead S.K.,
RA Carlow C.K.S., Jannasch H.;
RT "Intervening sequences in an Archaea DNA polymerase gene";
RL Proc. Natl. Acad. Sci. U.S.A. 89:5577-5581(1992).
RN [2]
RP PROTEIN SPLICING.
RX MEDLINE=93117083; PubMed=1475179;
RA Hodges R.A., Perler F.B., Noren C.J., Jack W.E.;
RT "Protein splicing removes intervening sequences in an archaea DNA
RL polymerase";
RN Nucleic Acids Res. 20:6153-6157(1992).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3', TO 5' EXONUCLEASE ACTIVITY.
CC -1- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
CC INTEIN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
CC ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
CC INTEIN.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -1- PFM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION
CC (INTEINS) FOLLOWED BY PEPTIDE LIGATION.
CC -1- BIOTECHNOLOGY: Used in the PCR method because of its high
CC thermostability and low error rate. Sold by New England Biolabs.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M74198; AAA72100.1; -;
CC EMBL; M74198; AAA72101.1; -;
CC DR PIR; S42459; S42459.
CC DR HSP; P56689; IIGO.
CC DR REBASE; 2621; PI-Tlii.
CC DR REBASE; 2621; PI-Tlii.
CC DR InterPro; IPR006172; DNA_pol_B.
CC DR InterPro; IPR006134; DNA_pol_B_dom.
CC DR InterPro; IPR006133; DNA_pol_B_exo.
CC DR InterPro; IPR003586; Hedgehog_hintC.
CC DR InterPro; IPR003587; Hedgehog_hintN.
CC DR InterPro; IPR006141; InteIn.
CC DR InterPro; IPR006142; InteIn.
CC DR InterPro; IPR004042; InteIn_endonuc.
CC DR InterPro; IPR004578; Pol2.
CC DR Pfam; PF00136; DNA_pol_B; 3.
CC DR Pfam; PF00104; DNA_pol_B_exo; 1.
CC DR PRINTS; PR00379; INTEIN.
CC DR SMART; SMO0305; HintC; 2.
CC DR SMART; SMO0306; HintN; 2.

DR SMART; SMO0486; POLBc; 1.
DR TIGRFAMS; TIGR01443; intein_Cterm; 2.
DR TIGRFAMS; TIGR01445; intein_Nterm; 2.
DR TIGRFAMS; TIGR00592; pol2; 2.
DR PROSITE; PS00116; DNA_POLYMERASE_B; FALSE_NEG.
DR PROSITE; PS00818; INTEIN_C_TER; 2.
DR PROSITE; PS00819; INTEIN_ENDONUCLEASE; 2.
DR PROSITE; PS00817; INTEIN_N_TER; 2.
KW Transferrase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
KW Protein splicing; Intrin homing.
FT CHAIN 1 494 DNA POLYMERASE, 1ST PART.
FT CHAIN 495 1032 ENDONUCLEASE PI-Tli I.
FT CHAIN 1033 1081 DNA POLYMERASE, 2ND PART.
FT CHAIN 1082 1471 ENDONUCLEASE PI-Tli I.
FT CHAIN 1472 1702 DNA POLYMERASE, 3RD PART.
SQ SEQUENCE 1702 AA; 197293 MW; 21D6B98C75F53B20 CRC64;
Query Match 83.4%; Score 3356; DB 1; Length 1702;
Best Local Similarity 42.7%; Pred. No. 8.7e-175;
Matches 726; Conservative 33; Mismatches 15; Indels 928; Gaps 2;
QY 1 MIFDTYITKDGKPIIRIFKKGEGFKIELDPHFQPIYVALLKDDSAIDBIKAIKGRHG 60
DB 1 MILDTYITKDGKPIIRIFKKGEGFKIELDPHFQPIYVALLKDDSAIEIKAIKGRHG 60
QY 61 KIVVDAVKVKKFKLGRDVEVWKLIEHPQDPALRGKIREHPAVIDIYEDIPFAKRY 120
DB 61 KTVRLDAVKRKFLGRDVEVWKLIEHPQDPVPMRGKIREHPAVVYDIYEDIPFAKRY 120
QY 121 LIDKGLIPMEGDEBELKMAPDIETFYHEGDFGKEIIMISYADEEERARVITWKNIDLPY 180
DB 121 LIDKGLIPMEGDEBELKMAPDIETFYHEGDFGKEIIMISYADEEERARVITWKNIDLPY 180
QY 181 VDVVSNREMIKRFQVVRKQDPDLITYGNDPDLPIYIKRAEKLGVTLGLGRDKEHPE 240
DB 181 VDVVSNREMIKRFQVVRKQDPDLITYGNDPDLPIYIKRAEKLGVTLGLGRDKEHPE 240
QY 241 PKIHRMGDSFAVEIKGRIHFDLPVVRTINLPYTLVAVVEAVLGTGKLGAEETAAI 300
DB 241 PKIHRMGDSFAVEIKGRIHFDLPVVRTINLPYTLVAVVEAVLGTGKLGAEETAAI 300
QY 301 WETESMKKLAQYSMEDARATYELGKFFPMEAEAKLIGQSVMDVSRSTGNLVEWYLL 360
DB 301 WETESMKKLAQYSMEDARATYELGKFFPMEAEAKLIGQSVMDVSRSTGNLVEWYLL 360
QY 361 RVAVERNELAPNKDEEYERRRRTTYLGGVKEPEKGLWENITYLDFCLYPSIIVTHN 420
DB 361 RVAVARNELAPNKDEEYERRRRTTYLGGVKEPEKGLWENITYLDFCLYPSIIVTHN 420
QY 421 VSPDTLEREGCKNDVAPIVGKFKDFPGFIPSIILGELITMRQEIKKQKATIDPIEKK 480
DB 421 VSPDTLEKGGCKNDVAPIVGKFKDFPGFIPSIILGELITMRQEIKKQKATIDPIEKK 480
QY 481 MLDYRQRAVKLHA----- 493
DB 481 MLDYRQRAIKLLANSILPNEWLPPIENGKIFVKIGFINSYMEKQENKVTVENTEVL 540
QY 494----- 493
DB 541 VNNLFAFSFNKKIKESVKKVKKALIRHKYKGAIEIQLSSGRKINITAGHSLFTVRNGEI 600
QY 494----- 493
DB 601 KEVSGDGIGKGLIIVAPKKIKLNEKGVSNIPELISDLSEETADIVMTISAKGRKNFFK 660
QY 494----- 493
DB 661 GMLATLRWFMGEENRIRTFNRYLFHLEKGLIKLPLRGYEVTDWELKKYKQLYEKL 720
QY 494----- 493
DB 721 SVKYNKREYLVNMFNBKIDPISYFPQKELEENKIGTNGFRNTNCLKVDEDFEKLGGY 780

```
QY 494 ----- 493
Db 781 VSEGAGAKNKTGIGSYSVKLYNEDPNVLSMKQNAEKFFGKVRVDRNCVSIKQMAYL 840
QY 494 ----- 493
Db 841 VMKCLGALAENKRIPSVILTSPEVPRVSFLEAYFTGDDIHPSKRFSLSTKSELLANQL 900
QY 494 ----- 493
Db 901 VFLNLSGSSVKIGFDSGVYRVINEDLPQTSRKNNTYYSNLPKEILRDVFGKEFQ 960
QY 494 ----- 493
Db 961 KMTWFKKELVDSKLNREKAKLEFFINGDIVLDVYKSKYKDYGVYVYDLSVEDNEN 1020
QY 494 ----- 543
Db 1021 FLVGFGLYAHNSYYGYGPKARWYSKCAESVTANGRHVYIEMTIRIEKFGFKVLYA 1080
QY 543 D----- 543
Db 1081 DSVSSESELIIRONKIRFVKIKDLFSKVDYSIGKEYCILEGVEALTLDDGKLWKPV 1140
QY 544 ----- 543
Db 1141 PYVHRHRANKRMFRIWLTNSWYIDVTEHSLIGLNTSKTKAKKIGERLKEVKPPELIGK 1200
QY 544 ----- 543
Db 1201 AVKSLICPNAPLKDENTKTSEIAVFWELVLVGDGNWGGDSRWAEYVYLGSLGKQABE 1260
QY 544 ----- 543
Db 1261 IKQKLEPLKTYGVISNYYPKNEKGFNLAKSLVKFMKRHPKDEKGRRKIPFMYELPV 1320
QY 544 ----- 543
Db 1321 TYIEAFLRGLFSADGTVTIRKGVPEIRLTNDADFLREVRLKLWIVGISNIFAEPTNR 1380
QY 544 ----- 543
Db 1381 YNGVSTGYSKHLRIKKNRFAERIGFLIERKQKLEHLKSARVKRNTIDFGDLVHVK 1440
QY 544 ----- 572
Db 1441 KYEPIPYGVYVDIEVEETHRPFANNILVHNTDGFYATIPGKPELIIKKAKKEFLNYNS 1500
QY 573 KLPGLLELEYEGFYLRGFFVAKRYAVIDEGRITTRGLEVVVRDMSIAKETQAKVLEA 632
Db 1501 KLPGLLELEYEGFYLRGFFVTKRYAVIDEGRITTRGLEVVVRDMSIAKETQAKVLEA 1560
QY 633 ILKEDSVEKAVIVKDVVEEIAKYQVPLEKLVHIOITKDLSEYKAIQPHVAIAKRLAAK 692
Db 1561 ILKEGSVEKAVEVRDWEKIAKYRVPLEKLVHIOITKDLSEYKAIQPHVAIAKRLAAK 1620
QY 693 GIKVRPTIISIVLKGSGKISDRVILLSEYQPKHKYDPDYVYIENQVLPVLRILEAFG 752
Db 1621 GIKVRPTIISIVLKGSGKISDRVILLTEYDPRKHKYPDYVYIENQVLPVLRILEAFG 1680
QY 753 YRKEDLKYSKQVGLDAILKK 774
Db 1681 YRKEDLYQSSKQTGLDAWLKR 1702
```

RESULT 3

```
DPOL_THEGO STANDARD; PRT; 773 AA.
AC PS6689;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

```
DE DNA polymerase (EC 2.7.7.7) (TO POL).
GN POL OR POLA.
OS Thermococcus gorgonarius.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=71997;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=99199230; PubMed=10097083;
RA Hopfner K.P., Eichinger A., Engh R.A., Laue F., Ankenbauer W.,
RA Huber R., Angerer B.;
RT "Crystal structure of a thermostable type B DNA polymerase from
RT Thermococcus gorgonarius.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:3600-3605(1999).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC PDB; 1TGO; 22-MAR-99.
DR InterPro; IPR006172; DNA_pol_B.
DR InterPro; IPR006134; DNA_pol_B_dom.
DR InterPro; IPR006133; DNA_pol_B_exo.
DR InterPro; IPR004578; Pol2.
DR Pfam; PF00136; DNA_pol_B_1.
DR Pfam; PF03104; DNA_pol_B_exo; 1.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLBc; 1.
DR TIGRFAMs; TIGR00592; pol2; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Hydrolase; Nuclease; Exonuclease; Multifunctional enzyme;
KW 3D-structure.
FT DISULFID 428 442
FT STRAND 2 10
FT TURN 11 12
FT STRAND 13 22
FT TURN 23 24
FT STRAND 25 31
FT STRAND 37 42
FT HELIX 45 47
FT HELIX 48 51
FT TURN 52 53
FT STRAND 55 58
FT TURN 59 60
FT STRAND 61 64
FT STRAND 67 75
FT TURN 76 77
FT STRAND 78 86
FT TURN 90 91
FT HELIX 92 102
FT TURN 104 105
FT STRAND 106 110
FT HELIX 116 123
FT TURN 124 125
FT STRAND 137 144
FT TURN 152 153
FT STRAND 157 164
FT TURN 165 166
FT STRAND 167 172
FT TURN 179 180
FT STRAND 181 183
FT HELIX 187 201
FT STRAND 205 208
FT HELIX 211 213
FT TURN 214 214
FT TURN 215 225
FT TURN 226 227
FT TURN 234 235
FT STRAND 240 243
FT STRAND 248 251
FT TURN 253 254
```


DR Pfam: PF03104; DNA pol B_exo; 1.
DR PRINTS; P00106; DNAPOLB.
DR SMART; SMO0486; POLBC; 1.
DR TIGRFAMS; TIGR00592; pol12; 1.
DR PROSITE; PS00116; DNA POLYMERASE B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Complete proteome.
FT VARIANT 532 532 R -> S (IN STRAIN GE23).
FT VARIANT 553 554 HE -> PN (IN STRAIN GE23).
FT VARIANT 756 756 K -> R (IN STRAINS GE23 AND ST 855).
FT CONFLICT 263 263 V -> A (IN REF. 1).
FT CONFLICT 277 277 A -> T (IN REF. 1).
FT CONFLICT 281 281 F -> V (IN REF. 1).
FT CONFLICT 320 320 A -> S (IN REF. 1).
FT CONFLICT 339 339 Q -> H (IN REF. 1).
FT CONFLICT 359 359 R -> T (IN REF. 1).
FT CONFLICT 391 391 K -> N (IN REF. 1).
SQ SEQUENCE 771 AA; 89496 MW; 110A87045A8A5522 CRC64;

Query Match 78.6%; Score 3164.5; DB 1; Length 771;
Best Local Similarity 76.1%; Pred. No 88-165;
Matches 588; Conservative 89; Mismatches 93; Indels 3; Gaps 2;

QY 1 MFTDTYITKQPIIRIFKNGKFKIELDPHPQYIYALLKDSADIEKAIGERHG 60
DB 1 MIIDADYITDQKPIIRIFKNGKFKVEYDRTFRPIYIYALLKDSADIEKITAERHG 60

QY 61 KIVRVVDVAVKKKELGRDVEYVKKLIFEHPQDVPALRGKIREHPAVIYIYEDIPAKRY 120
DB 61 KIVRITEVEKVKKELGRPIEVKKLYLEHPQDVPALRGKIREHPAVIYIYEDIPAKRY 120

QY 121 LIDKGLIPMEGDEELKLMADFIETFYHGDGFGKGEIIMISYADDEEAEARVITKNIDLPY 180
DB 121 LIDKGLTPMEGNEELTFIYDITLYHEGEEFGKGIIMISYADDEGAKVITKSIDLPY 180

QY 181 VDVVSNRERMIKRFQIVREKDPDLITNGDNFPLPKYIKRAEKLGVTLILGRKHEPE 240
DB 181 VEVVSEERMIKRLVIREKDPDLITNGDNFPLPKYIKRAEKLGVTLILGRD--NSE 238

QY 241 PKIHRMGDSFAVEIKRIFHDFLPVVRRTINLPTVLEAVYEAVALGKTKSLGAEIAAI 300
DB 239 PKMQRMGDSLAVELKRIHDFLPVVRRTINLPTVLEAVYEAIFGSKSKVYAEIAEA 298

QY 301 WETESMKLAQSMEDATATVGLKEFPPEAEALKLGQSVWDSVSSSTGNLVEWILL 360
DB 299 WETGKLERVAKYSMEDAKVTTELQKEFPPEAEALRLVGQVPWDSVSSSTGNLVEWFL 358

QY 361 RVAYERNELAPKPDDEEYRRRLRTTYLGGYVKEPERGLWENITVLDPRCLYPSIIVTHN 420
DB 359 RXAYERNELAPKPDEREYERLRRESYEGYVKEPEKGLWEGIVSLDFRSLYPSIIVTHN 418

QY 421 VSPDTLRECKYDVAPIVGKPKDPGPTPSILGELITNWRQIKKKKATIDPIKK 480
DB 419 VSPDTLNRECKEYDVAPQVGHFRFCDFGPTPSILGNLLEERQIKKRMKESKDPVEKK 478

QY 481 MLDYRQAVKLHANSYGYMGYPKARWYSKECAESVTAGRHYIEMTIKEEKGFKVL 540
DB 479 LLDYRQRAIKILANSYGYGYGAKARWYCKECAESVTAGRQYIDLVRRELESR-GFKVL 537

QY 541 YADTGFYATIGEPETIKKAKFLKYINSLPGLLEVEYGYLRFVFKRYAVI 600
DB 538 YTDGLYATIPGAKHEEIKKALFVEYINSLPGLLEVEYGYARGFVTKKYALI 597

QY 601 DEEGRTTTRGLVWRDWESEIAKETQAKVLEALKEDSVKAVEIVKDWVEIAKQVPL 660
DB 598 DEEGKIVTGLVWRDWESEIAKETQAKVLEALKGNVDEAKVKEVTEKLSYIEIPP 657

QY 661 EKLVIHQITKDSLEYKATGPHVAIAKLAAGKIKVRPGTIISYIVLRGSKISDRVILL 720
DB 658 EKLVIYEQITRPSLEYKATGPHVAIAKLAAGKIKVKGWIGYIVLRAGDGPISKRAI 717

QY 721 SEYDPKKHKYDPDYIENQVLPVLRILLEAFGYRKEDLKYQSSKQVGLDAWLK 773

Db 718 EEPDPKKHKYDAEYIENQVLPVLRILAFGYRKEDLKYQTKQVGLGAWLK 770

RESULT 6
ID DPOL_PYRPU STANDARD; PRT; 775 AA.
AC P80061; P95584;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) (Pfu polymerase).
GN POL OR PF0212.
OS Pyrococcus furiosus, and
OS Pyrococcus woesei.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261, 2262;
RN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RP SPECIES=P.furiosus; STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422;
RA MEDLINE=93181200; PubMed=8441634;
RA Uemori T., Ishino Y., Toh H., Asada F., Kato I.;
RT "Organization and nucleotide sequence of the DNA polymerase gene from
the archaean Pyrococcus furiosus.";
RL Nucleic Acids Res. 21:259-265(1993).
RN [2] SEQUENCE FROM N.A.
RP SPECIES=P.furiosus; STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3] SEQUENCE OF 396-418 FROM N.A.
RP SPECIES=P.furiosus;
RA MEDLINE=92107689; PubMed=1762925;
RA Mathur E.J., Adams M.W.W., Callen W.N., Cline J.M.;
RT "The DNA polymerase gene from the hyperthermophilic marine
archaeobacterium, Pyrococcus furiosus, shows sequence homology with
alpha-like DNA polymerases.";
RL Nucleic Acids Res. 19:6952-6952(1991).
RN [4] SEQUENCE OF 396-502 FROM N.A.
RP SPECIES=P.furiosus;
RA Mathur E.J., Adams M.W.W., Callen W.N., Cline J.M.;
RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
RN [5] SEQUENCE FROM N.A.
RP SPECIES=P.woesei; STRAIN=DSM 3773;
RA MEDLINE=98432910; PubMed=9758761;
RA Dabrowski S., Kur J.;
RT "Cloning and expression in Escherichia coli of the recombinant his-
tagged DNA polymerases from Pyrococcus furiosus and Pyrococcus
woesei.";
RL Protein Expr. Purif. 14:131-138(1998).
RN [6] SIMILARITY TO OTHER POLYMERASES.
RA MEDLINE=92253396; PubMed=1579479;
RA Forterre P.;
RT "The DNA polymerase from the archaeobacterium Pyrococcus furiosus does
not testify for a specific relationship between archaeobacteria and
eukaryotes.";
RL Nucleic Acids Res. 20:1811-1811(1992).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA}[N].
CC -1- SUBUNIT: Monomer.
CC -1- BIOTECHNOLOGY: Because pfu DNA polymerase exhibits the lowest
error rate of any thermostable DNA polymerase studied, it is
routinely used for PCR. It is sold by Promega.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

DR	PRINTS; PRO0106; DNAPOLB.	
DR	PRINTS; PRO0379; INTEN.	
DR	SMART; SM00305; HintC; 1.	
DR	SMART; SM00306; HintN; 1.	
DR	SMART; SM00486; POLBc; 1.	
DR	TIGRFAMS; TIGR01443; intein_Cterm; 1.	
DR	TIGRFAMS; TIGR01445; intein_Nterm; 1.	
DR	TIGRFAMS; TIGR00592; POL2; 1.	
DR	PROSITE; PS00116; DNA POLYMERASE B; 1.	
DR	PROSITE; PS00818; INTENIN_C_TER; 1.	
DR	PROSITE; PS00819; INTENIN_ENDONUCLEASE; 1.	
DR	PROSITE; PS00817; INTENIN_N_TER; 1.	
KW	Transferase; DNA-directed DNA polymerase; DNA replication;	
KW	DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;	
KW	Protein splicing; Intronic homing.	
FT	CHAIN 1 492 DNA POLYMERASE, 1ST PART.	
FT	CHAIN 493 1029 ENONUCLEASE PI-PSPI.	
FT	CHAIN 1030 1312 DNA POLYMERASE, 2ND PART.	
SQ	SEQUENCE 1312 AA; 152852 MW; B62518805641D26A CRC64;	
Query Match 72.9%; Score 2934.5; DB 1; Length 1312;		
Best Local Similarity 45.6%; Pred. No. 5e-152;		
Matches 597; Conservative 86; Mismatches 87; Indels 539; Gaps 3;		
QY	1 MIFDTYITKDKPIIRIPKENGSEFKIELDPHFQPIYALLKDDSAIDKAIKGERHG 60	
DB	1 MILDADYITDCKPIIRIPKENGSEFKVEYDRNFRPIYALLKDDSQIDEVRKITAERHG 60	
QY	61 KIVRVVDAYKKKGLGRDVEYWKLIFFHPQDPVPAIRKIRHPAVIDIYEDIPFAKRY 120	
DB	61 KIVRIIDAEKVRKGLGRDIEVWRLYFHPQDPVPAIRKIRSHSAVIDIYEDIPFAKRY 120	
QY	121 LIDKGLIPMEGDEELKMAFDIETVYHEGDEFGKEIIMISYADDEEARVITWKNDLPY 180	
DB	121 LIDKGLIPMEGDEELKMAFDIETLYHEGEERAKGPIIMISYADDEEARVITWKNDLPY 180	
QY	181 VDVNSNERMIRFQIVREKDPDVLITVNGNFDLPYLIRKAEKLGVTLLGRDKHEPE 240	
DB	181 VEWSSEREMIRFLKVRREKDPDVIITVNGSDFDLPYLVRKAEKLGKILPLGRDGS--E 238	
QY	241 PKIHRMGDSFAYEIKGRHFDLPVVRRTINLPTVTLAATVAVLGKTKSLGABEIAAI 300	
DB	239 PKWQLRGDMTAVEIKGRHFDLYHVRRTINLPTVTLAATVAVLGKTKSLGABEIAEA 298	
QY	301 WTESMKKLAQYSMEDAPATVELKGFPPMPDAELAKIGOSVWDVSRSSSTGNLWVYLL 360	
DB	299 WETGKGLERVAKYSMEDAKVTVELGREFFPMEQSLRVGQPLWDVSRSSSTGNLWVYLL 358	
QY	361 RVAYERNELAPKPDDEEYRRLRTLYGGYVKEBERGLWENITVLDPRCLYPSIIVTHN 420	
DB	359 RXAYERNELAPKPDDEEYRRLRYSAGYVKEBERGLWGLVSLDPRSLYPSIIVTHN 418	
QY	421 VSPDTLREGCKNYDVAPITVGYKFCDFPGFIPSPILGELITWQRIKKMKATIDPIEKK 480	
DB	419 VSPDTLNRGCREYDVAPVGVGHKFCDFPGFIPSPILKRLDLDERQEIKKMKASKDPIEKK 478	
QY	481 MLDYRQRAVK----- 490	
DB	479 MLDYRQRAIKILANSILPEEVPFLIKNGKVIKIFRIGDFVDGLMKANQKVKKTGTDEVLE 538	
QY	491 ----LHA----- 493	
DB	539 VAGIHAFSFRDSSKKARVMKAVIRHRYSGNVYRIVLNSGRKIITTEGHSIFVYVRNDL 598	
QY	494 ----- 493	
DB	599 VEATGEDVKIGDLLAVPRSVNLPEKRLNIVELLNLSPETEDIIITIPVKGRKNPFK 658	
QY	494 ----- 493	
DB	659 GMLRTLWIFGBEKRVRTASVRLHLENLYIRLKIIGVDIIDKGLKRYTLVEKLVDV 718	
QY	494 ----- 493	

RESULT 8		
DPOL_PYRHO	STANDARD;	PRT; 1235 AA.
AC	OS9610;	
DT	15-DEC-1998 (Rel. 37, Created)	
DT	15-DEC-1998 (Rel. 37, Last sequence update)	
DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	DNA polymerase (EC 2.7.7.7) [Contains: Pho pol intein (Pho Pol I	
DE	intein)]	
GN	POL OR PH1947 OR PHB047.	
OS	Pyrococcus horikoshii.	
OC	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;	
OX	Pyrococcus.	
NCBI	TaxID=53953;	
EN	[1]	
EP	SEQUENCE FROM N.A.	
RC	STRAIN=OT3;	
RX	MEDLINE=98344137; PubMed=9679194;	
RA	Kawarayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,	
RA	Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,	
RA	Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,	
RA	Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A.,	
RA	Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,	
RA	Masuchi Y., Shizuya H., Kikuchi H.;	
RT	"Complete sequence and gene organization of the genome of a hyper-	
RT	thermophilic archaeobacterium, Pyrococcus horikoshii OT3.,"	
RL	DNA Res. 5:55-76(1998).	
CC	-!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate	
CC	+ {DNA}(N).	
CC	-!- PIM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES	
CC	A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)	
CC	FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).	
CC	-!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.	

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP000007; BAA31074.1; --
CC PIR; C71210; C71210.
CC HSP; P56689; ITGO.
DR InterPro; IPR006172; DNA_pol_B.
DR InterPro; IPR006134; DNA_pol_B_dom.
DR InterPro; IPR006133; DNA_pol_B_exo.
DR InterPro; IPR003586; Hedgehog_HintC.
DR InterPro; IPR003587; Hedgehog_HintN.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR006142; Intein.
DR InterPro; IPR004042; Intein_endonuc.
DR InterPro; IPR004578; Pol2.
DR Pfam; PF00136; DNA_pol_B; 1.
DR Pfam; PF03104; DNA_pol_B_exo; 1.
DR PRINTS; PR00106; DNA_POLB.
DR PRINTS; PR00379; INTEIN.
DR SMART; SM00305; HintC; 1.
DR SMART; SM00306; HintN; 1.
DR SMART; SM00486; POLBc; 1.
DR TIGRFAMs; TIGR01443; Intein_Cterm; 1.
DR TIGRFAMs; TIGR01445; Intein_Nterm; 1.
DR TIGRFAMs; TIGR00592; pol2; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
DR PROSITE; PS00818; INTNIN_CTER; 1.
DR PROSITE; PS00819; INTNIN_ENDONUCLEASE; 1.
DR PROSITE; PS00817; INTNIN_NTER; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Nucleocatalytic cleavage; Protein splicing;
KW Complete proteome.
FT CHAIN 1 492 DNA POLYMERASE, 1ST PART (POTENTIAL).
FT CHAIN 493 952 PHO POL INTEIN (POTENTIAL).
FT CHAIN 953 1235 DNA POLYMERASE, 2ND PART (POTENTIAL).
SQ SEQUENCE 1235 AA; 143086 MW; 73CC7AA14873CCE4 CRC64;

Query Match 71.7%; Score 2886; DB 1; Length 1235;
Best Local Similarity 46.5%; Pred. No. 2e-149;
Matches 574; Conservative 107; Mismatches 89; Indels 464; Gaps 5;

QY 1 MIFDTYITKDGPIIRIPFKKENGSEKIFELDPHPQPIYVALLKDDSAIDEIKAKGERHG 60
DB 1 MILDADYITEDGPIIRIPFKKENGSEKIFELDPHPQPIYVALLKDDSAIDEIKAKGERHG 60

QY 61 KIVRVVDVAVKXKFLGRDVEVVKLIFHPQDVPALRGKIRHPAVIDIYEDIPFAKRY 120
DB 61 KVRIVETSKIQKFLGRFIEVVKVLEHPQDVPALRGKIRHPAVIDIYEDIPFAKRY 120

QY 121 LIDKGLIPMEGDEBELKMAFDIETPFYHGEDEFGKBEIMISYADDEEAAVITWKNIDLPY 180
DB 121 LIDKGLTPMEGNEKLTFLAVDIETLYHEGEFEFGKGPVIMISYADDEEAAVITWKNIDLPY 180

QY 181 VDVVSNEREMIKFPVQIVAREKDPDLITVNGDNFDLPYLTKAEKLGVTLLGRKPERPE 240
DB 181 VEVVSNEREMIKLIRVIREKDPDLITVNGDNFDLPYLTKAEKLGVTLLGRD--NSE 238

QY 241 PKIHRGDSFAVIEKGRHIFDLFPVVRRTINLTPTYLEAVYEAVLTKTSKAGAEIAAI 300
DB 239 PKQKMGDSLAVIEKGRHIFDLFPVVRRTINLTPTYLEAVYEAVLTKTSKAGAEIAAI 298

QY 301 WETEESKKLAQVSMEDARATYELGKFFPMEAEKLAGKOSVWDYSRSTGNLVEWYLL 360
DB 299 WETGEGLERVAKVSMEDAKVYELGKFFPMEAEKLAGKOSVWDYSRSTGNLVEWYLL 358

QY 361 RVAVERNELAPNKDEEYERRRLRTTYLGGYVKEPERGLWENTTYLDFRLYPSIIVTNN 420
DB 361 RVAVERNELAPNKDEEYERRRLRTTYLGGYVKEPERGLWENTTYLDFRLYPSIIVTNN 418

DB 359 RKAVERNELAPNKDEEYERRRLRESYEGYVKEPKLWEGIVSLDFRSLYSIIITNN 418
QY 421 VSPDTEREGCKNDVAPIVGYKFCDFPGFIPSLIGELITMRQETKQKWKATIDPIEKK 480
DB 419 VSPDTLANEGCEEYDVAPKVGHRFCDFPGFIPSLIGELITMRQETKQKWKATIDPIEKK 478
QY 481 MLDYRQRAVKL----- 491
DB 479 LLDYRQRAIKILANSILPDEWLPIVENEKRVFKVIGDFIDRETEENAEVVKRDEGTEILE 538
QY 492 ----- 491
DB 539 VKDLKALSFNRETCKSELKVKALIRHRYSGKYSIKLKGRRIKITSGHSLPSVQNGKL 598
QY 492 ----- 491
DB 599 VKVRGDELKPGDLVVPGRKLPEKQVNLVELLLKLPEBETSNIWMTIPVKGRKNFFK 658
QY 492 ----- 491
DB 659 GMLKTLVYFGEGERPRTAGRYLXHLERLGYVVKLRGCEVLDWESLKRVRKLYETLIKN 718
QY 492 ----- 491
DB 719 LKYNCSRAYMVFENSLRDVSLMPIELKEWIIICEPRGPKIGTIFDVDDSFALGLYII 778
QY 492 ----- 493
DB 779 SSGDVEKORVPHSKDQNVLEDAKLAELFKGVRRGRGYIEVSGKISHAIFRVLAEGRK 838
QY 494 ----- 493
DB 839 IPEFTFTSPMDIKVAPLKGNGNABELAFSTKSELLNQLILLNLSIGVSDIKIEHEKGV 898
QY 494 ----- 500
DB 899 RYVYINKESSNGDIVLDSVESIEVEKGYVYVDSLEDNENFLVGFGLLYAHNSYGY 958
QY 501 GYPMARWYKECAESVTAMGRHYIEMTKIEBKFGKVLVADTGFATITCEKP-EPI 559
DB 959 GYKASWYKECAESVTAMGRQYIDLVRRLEAR-GFKVLYITDGLYATIPGVKDWEEV 1017
QY 560 KKAKKFLKXVNSKLPGLLEVEGFGYLRGFPVAKRVAVIDEGRITTRGLEVVRDWS 619
DB 1018 KRSALEFVYVNSKLPGLLEVEGFGYLRGFPVAKRVAVIDEGRITTRGLEVVRDWS 1077
QY 620 EIAKETQAKVLEAILKEDSVKAVEIVKDVVEIAKYQVPLEKLVHBOITKDLSEYKAI 679
DB 1078 EIAKETQARVLEAILKHGNEVBAKIVKDVTEKLTNYEVPPEKLVYEQITRPINEYKAI 1137
QY 680 GPHVATAKRLAAGIKVRPCTIISYIVLRSGKISDRVILLSEYDPKHKYDPDYIENQ 739
DB 1138 GPHVAVAKRLMARGIKVKGPMVIGYIVLRGDDGPIKRAISIEFDFPRKHKYDAEYIENQ 1197
QY 740 VLPVAVRILEAFGRYKEDLKYQSSKQVGLDANLK 773
DB 1198 VLPVAVRILEAFGRYKEDLKYQSSKQVGLDANLK 1231

RESULT 9
ID DPOL THEFM STANDARD; PRT; 1523 AA.
AC 174918;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) (Pol Tfu) [Contains: Endonuclease PI-TfuI
DE (EC 3.1.-.-) (Tfu pol-1 intein); Endonuclease PI-TfuI (EC 3.1.-.-)
DE (Tfu pol-2 intein)].
GN POL.
OS Thermococcus funicolans.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.

OX NCBI_TaxID=46540;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ST557;
RA Camdon M., Querellou J.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION OF INTENS.
RC STRAIN-ST557;
RX MEDLINE=20112788; PubMed=10644683;
RA Saves I., Ozanne V., Dietrich J., Maason J.-M.;
RT "Intens of Thermococcus fumicolans DNA polymerase are endonucleases
RT with distinct enzymatic behaviors.";
RL J. Biol. Chem. 275:2335-2341 (2000).
CC -!- FUNCTION: PI-Tfui recognizes and cleaves a minimal sequence of 16
CC base pairs (bp) on supercoiled DNA with either Mn(2+) or Mg(2+) as
CC cofactor. It cleaves linear DNA only with Mn(2+) and requires a
CC 19-bp minimal recognition sequence. The optimal temperature for
CC activity is 70 degrees Celsius.
CC -!- FUNCTION: PI-Tfui is a highly active homing endonuclease using
CC Mg(2+) as cofactor. Its minimal recognition and cleavage site is
CC 21 bp long either on linear or circular DNA substrates. Its
CC endonuclease activity is strongly inhibited by the 3' digestion
CC product, which remains bound to the enzyme after the cleavage
CC reaction. The optimal temperature for activity is 70 degrees
CC Celsius.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION
CC (INTENS) FOLLOWED BY PEPTIDE LIGATION.
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -!- SIMILARITY: IN THE INTEN SECTION; BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 269882; CAA93738.1; -
CC HSP; P56689; ITGO.
CC REBASE; 4500; PI-Tfui.
CC REBASE; 4501; PI-Tfui.
CC InterPro; IPR006172; DNA pol B.
CC InterPro; IPR006134; DNA pol B_dom.
CC InterPro; IPR006133; DNA pol B_exo.
CC InterPro; IPR003586; Hedgehog hintC.
CC InterPro; IPR003587; Hedgehog hintN.
CC InterPro; IPR006141; Intein.
CC InterPro; IPR006142; INTEIN.
CC InterPro; IPR004042; Intein_endonuc.
CC InterPro; IPR004578; Pol2.
CC Pfam; PF00136; DNA_pol_B; 2.
CC Pfam; PF03104; DNA_pol_B_exo; 1.
CC PRINTS; PR00379; INTEIN.
CC SMART; SM00305; HintC; 2.
CC SMART; SM00306; HintN; 2.
CC SMART; SM00486; POLB; 1.
CC TIGRfams; TIGR01443; intein_Cterm; 2.
CC TIGRfams; TIGR01445; intein_Nterm; 2.
CC TIGRfams; TIGR00592; pol2; 1.
CC PROSITE; PS00116; DNA_POLYMERASE_B; FALSE_NEG.
CC PROSITE; PS00818; INTEIN_C_TER; 2.
CC PROSITE; PS00819; INTEIN_ENDONUCLEASE; 2.
CC PROSITE; PS00817; INTEIN_N_TER; 2.
CC Transferrase; DNA-directed DNA polymerase; DNA replication;
CC DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
CC Protein splicing; Intronic homing; Magnesium; Manganese.
CC CHAIN 1 406
CC

FT	CHAIN	407	766	ENDONUCLEASE PI-TFUI.
FT	CHAIN	767	900	DNA POLYMERASE, 2ND PART.
FT	CHAIN	901	1282	ENDONUCLEASE PI-TFUI.
FT	CHAIN	1283	1523	DNA POLYMERASE, 3RD PART.
SO	SEQUENCE	1523	AA; 175917	MW; 7A2AC236BF2E5F5 CRC64;

Query Match 69.8%; Score 2809; DB 1; Length 1523;
Best Local Similarity 39.0%; Pred. No. 4.1e-145;
Matches 593; Conservative 86; Mismatches 91; Indels 752; Gaps 4;

QY	1	MIPDTYITKDGKPIIRIKFNGEPEKIELDPHFQPIYIALLKDDSAIDEIKAKGERHG	60
DB	1	MILDTYITTEDGRFVIRVFKENGFEKIEYDRDFEYIYALLKDDSAIDEIKAKTASRHG	60
QY	61	KIVRVVDVAVKVKKFLGRDVEVWKLIFFHPQDPVPAIRKIRHPAVIDIYVDIPFAKY	120
DB	61	TTVRVVRAGVKVKFLGRVIEVWKLYFTHQDPVPAIRKIRHPAVIDIYVDIPFAKY	120
QY	121	LIDKGLIPMEGDEELKMAFDIETFYHGEDEFGKGEIIMISYADBEAEAVITWKNIDLPY	180
DB	121	LIDKGLIPMEGDEELKMAFDIETFYHGEDEFGKGEIIMISYADBEAEAVITWKNIDLPY	180
QY	181	VDVVSNEREMIKFVOIVREKDPDVLITVNGDNFDLPYLKRAEKLGVTLGRDKEHPE	240
DB	181	VDVVSNEKEMIKFVKVKEKDPDVLITVNGDNFDLPYLKRAEKLGVTLGRDKEHPE	240
QY	241	PKIHRMGDSFAVEIKRIHFDLPFVVRRTINLPTVLEAVYPAVLGKTSKLGAEIAAI	300
DB	239	PKIQRMGDRFAVEIKRIHFDLPVIRHTINLPTVLEAVYPAVLGKTSKLGAEIAAI	298
QY	301	WETEESSMKLAQYSEMDARATYELGKEFPFMAELAKLIGQSVWDSRSTGNLVEWYLL	360
DB	299	WETEGELRVARYSMEDAKVTYELGKEFPFMAELAKLIGQSVWDSRSTGNLVEWYLL	358
QY	361	RVAYERNELAPNKPDEEYRRLRTYLGKYKPERGLWENITVLDPRC	410
DB	359	RKAYERNELAPNKPSEGLERR-RGGYAGGYKPERGLWENIAYLDFRCHPADTKVIK	417
QY	411	-----	410
DB	418	GGGVNISEVREGDYVLGIDGKQKQVORVWEYDEGELVNNINGLKTNPKNLPPVARTERQ	477
QY	411	-----	410
DB	478	TAIRDSLAKSFLTKVKVGLITLPLFEKIGKIEREDVPEEILKGLAGIILAEGLLARK	537
QY	411	-----	410
DB	538	DVEYFDSRGGKRVSHQVREITVGAQBEDFORRIVYIPERLFVGTPTSVYRKKNTNATIF	597
QY	411	-----	410
DB	598	KVAKKEVYLVRREIMDGIENLHAPSLRGFFBGDGSVNVKRVTVVNVQGTNNEMKIEVVS	657
QY	411	-----	410
DB	658	KLLNKLGIPIHRYTYDYTEREKTMTTHILEIAGRDGLIFQITGVFISTEKNMALEAIR	717
QY	411	-----	410
DB	718	NREVNRLNNAFYTLADFTAKTEYKGVYDLTLEGTFFYFANGILITNSLYPSIIISHN	777
QY	421	VSPDTLEREGCKNVDVAPIGVYKFCDFPGFIPSLBELITWRQIKKKMKATIDPIBK	840
DB	778	VSPDTLNRGCGEYDEAPQVGHRCDFPGFIPSLGDLDERQKVKKMKATIDPIBK	837
QY	481	MLDYQRAVKLHANSYGYGMGYKARWTSKECAESVTAWGRHYIEMTKETEEKFGKVL	540
DB	838	LLDYQRAIKILANSFYGYGYKARWTSKECAESVTAWGRHYIEMTKETEEKFGKVL	897
QY	541	YAD-----	543
DB	898	YADSVTGDTEVTRIRNGRIEFVPIEKLPERVDHVRGEKEYCVLGGVEALTLDNRGLVWK	957

```

QY 544 ----- 543
Db 956 KVEYVNRHTDKRIYRVWFTNSWYLDVTDHSLIGYLNWTSVKVPGKPLKRLVEVRPEEL 1017
QY 544 ----- 543
Db 1018 GGVKSLITNPRPIARTIKANPIAVKLMELIGLVGDGNGWGSQSNWAKYVYVGLSCGLDKA 1077
QY 544 ----- 543
Db 1078 EIERKVLNPLREASVISNYDKSKGDVSVLSKMLAGFWKVFKDENGKKAIFSMFNLP 1137
QY 544 ----- 543
Db 1138 REYIEAFLRFSADCTVSLRRGIPEIRLTSVNRELSDAVRKLLWLVGVSNSLFTETKPN 1197
QY 544 ----- 543
Db 1198 RYLESSGTHSVIRIKNKRFPADRIGLIDRSTKLSNGLGHTNKKRAYKYDFDLVYP 1257
QY 544 ----- 571
Db 1258 RKIEETTYDGYVYDIEVEGTHRFANGILVHNTDGFATIPGADAETVKKAREFLNYIN 1317
QY 572 SKLPGLLELEYEGFYLRGFVAKRYAVIDEGRITTRGLEVVRDWSSEIAKETOAKVLE 631
Db 1318 PKLPGLLELEYEGFYRGRFVTKKYAVIDEGRITTRGLEIVREDHSEVAKETOQAVLE 1377
QY 632 AILKEDSVKAEIVKQVVEEIAKQVPLEKLVHIEQITKDLSEYKALGPHVAIAKRLAA 691
Db 1378 AILRHGDVSEAVRIVEKTEKLSKVEVPEKLVHIEQITRLEKDYKAGPHVAIAKRLAA 1437
QY 692 KGIKVRPGTHIISVILVSGSKISDRVILLSEYDPPKKHYDDYVYIENQVLPVAVLRILEAF 751
Db 1438 RGIVAPGVIVSVILVSGSGRIGDRITPFDEBDFPKHYDAEYVYIENQVLPVAVLRILKAF 1497
QY 752 GYKEDOLKYOSKOVGLDGLWK 773
Db 1498 GYKEDLRYQKRVGLGAWLK 1519

RESULT 10
DPOL PYRKO
ID DPOL PYRKO STANDARD; PRT; 1671 AA.
AC P77933;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) [Contains: Endonuclease P1-PkoI
DE (EC 3.1.1.-) (Pko pol-1 intein) (IVS-A); Endonuclease P1-PkoII
DE (EC 3.1.1.-) (Pko pol-2 intein) (IVS-B)].
GN POL.
OS Pyrococcus kodakaraensis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OX Thermococcus.
OC NCBI_TaxID=69014;
RN [1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=KOD1;
RX MEDLINE=98027387; PubMed=9361436;
RA Takagi M., Nishioka M., Kakiwara H., Kitabayashi M., Inoue H.,
RA Kawakami B., Oka M., Imanaka T.;
RT "Characterization of DNA polymerase from Pyrococcus sp. strain KOD1
RT and its application to PCR";
PL Appl. Environ. Microbiol. 63:4504-4510(1997).
RN [2]
CHARACTERIZATION OF INTEINS.
RC STRAIN=KOD1;
RX MEDLINE=98416198; PubMed=9742242;
RA Nishioka M., Fujiwara S., Takagi M., Imanaka T.;
RT "Characterization of two intein homing endonucleases encoded in the
RT DNA polymerase gene of Pyrococcus kodakaraensis strain KOD1.";
```

```

RL Nucleic Acids Res. 26:4409-4412(1998).
CC -!- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
CC -!- INTEIN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
CC -!- ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
CC -!- INTEIN. P1-PKOI RECOGNIZES 5'-GATTAGTACCTGTACC-3' AND P1-PKOII
CC -!- RECOGNIZES 5'-CAGTACTAGCTGTAC-3'. BOTH ARE THERMOSTABLE.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC -!- + {DNA}(N).
CC -!- P1M: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC -!- A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC -!- FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -!- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
CC -!- ENDONUCLEASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D29671; BAA06142.2; -.
CC PDB; 1GCX; 28-JAN-03.
CC REBASE; 3792; P1-PKOI.
CC REBASE; 3793; P1-PKOII.
CC InterPro; IPR006172; DNA_pol_B.
CC InterPro; IPR006134; DNA_pol_B_dom.
CC InterPro; IPR006133; DNA_pol_B_exo.
CC InterPro; IPR003586; Hedgehog_hintc.
CC InterPro; IPR003587; Hedgehog_hintn.
CC InterPro; IPR006141; Intein.
CC InterPro; IPR006142; INTEIN.
CC InterPro; IPR004042; Intein_endonuc.
CC InterPro; IPR004578; Pol2.
CC Pfam; PF00136; DNA_pol_B; 3.
CC Pfam; PF03104; DNA_pol_B_exo; 1.
CC PRINTS; PR00379; INTEIN.
CC SMART; SM00305; HintC; 2.
CC SMART; SM00306; HintN; 2.
CC SMART; SM00486; POLBc; 1.
CC TIGRFAMs; TIGR01443; intein_Cterm; 2.
CC TIGRFAMs; TIGR01445; intein_Nterm; 2.
CC TIGRFAMs; TIGR00592; pol2_2.
CC PROSITE; PS00115; DNA_POLYMERASE_B; 1.
CC PROSITE; PS00818; INTEIN_C_TERM; 2.
CC PROSITE; PS00819; INTEIN_N_TERM; 2.
CC PROSITE; PS00817; INTEIN_N_TERM; 2.
CC TRANSFERASE; DNA-directed DNA polymerase; DNA replication;
CC DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
CC Protein splicing; Intronic homing; 3D-structure.
CC CHAIN 1 406 DNA POLYMERASE, 1ST PART.
CC CHAIN 2 407 766 ENDONUCLEASE P1-PKO I.
CC CHAIN 3 767 851 DNA POLYMERASE, 2ND PART.
CC CHAIN 4 852 1388 ENDONUCLEASE P1-PKO II.
CC CHAIN 5 1389 1671 DNA POLYMERASE, 3RD PART.
CC SEQUENCE 1671 AA; 193490 MW; 4A17F1C9120EE455 CRC64;

Query Match 69.1%; Score 2782; DB 1; Length 1671;
Best Local Similarity 36.0%; Pred. No. 1.3e-143;
Matches 602; Conservative 82; Mismatches 86; Indels 900; Gaps 4;

QY 1 MIFDTYITDGGKPIIRFKKENGKFKILDPHFQPIYVALLKDDSAIDEIKAKGERHG 60
Db 1 MILDYITDGGKPIIRFKKENGKFKILDPHFQPIYVALLKDDSAIDEIKAKGERHG 60
QY 61 KIVRVDAVKKKFLGRDVEVWKLIFEPQDVPAALRGKIREHPAVIYIYEDIPFARY 120
Db 61 TVWTVKRVKKKFLGRDVEVWKLIFEPQDVPAALRGKIREHPAVIYIYEDIPFARY 120
QY 121 LIDKGLIPMGDDDELKMAFDITTFVHGDGFKGSIIMISVADEEAEARVITWKNIDLPY 180
Db 121 LIDKGLIPMGDDDELKMAFDITTFVHGDGFKGSIIMISVADEEAEARVITWKNIDLPY 180
```

Db 121 LIDKGLVPMGEDELKMLAFDIETLYHEGEBAEFGPILMSIYADSEGARVITMKNVDLPY 180
QY 181 VDVSNRERMIKRFQVIREKDPDVLITVNGDNFDPYLIKRAEKLGVTLILGRDKEHPE 240
Db 181 VDVSTEREMIKRFURVREKDPDVLITVNGDNFDPYLIKRAEKLGVTLILGRDKEHPE 238
QY 241 PKIHRMGDSFAVEIKRIFHDFLPVVRRTINLPTTLEAVYBAVLGKTKSLGAEIEAAI 300
Db 239 PKIQMGDRFAVEIKRIFHDFLPVVRRTINLPTTLEAVYBAVLGKTKSLGAEIEAAI 298
QY 301 WETESMKLAQYSMEDARATVELGKEFPFMAEALAKIGQSVWDVSSSTGNLVEWYLL 360
Db 299 WETGENLERVARYSMEDAKVYVELGKEFLPMAQSLRIGQSLWDSVSSSTGNLVEWYLL 358
QY 361 RVAYERNELAPNKPDEEYERRLRTYLGYYVKEPERGLWENITVLDPERC 410
Db 359 RYAYERNELAPNKPDEKELARR-RQSYEGGYKEPERGLWENITVLDPERCHPADTKVVK 417
QY 411 410
Db 418 KGKIINISEVQEGYVIGIDGQVRVKWWEYDKGELVNLNGLKCTPNHKLPPVTKNRQ 477
QY 411 410
Db 478 TRIRDSLAKSLTKVKKIIITPLPYELGRATSENIPEEVLKGLAGILLASGTILRK 537
QY 411 410
Db 538 DVEYFDSRRKRRIHQYRVEITIGKDEEPRDRITYIFERLFGITPISKKGTNAVTL 597
QY 411 410
Db 598 KVAKKNVYLKVEIMDNIESLHAPSVLRGFFGDSVNRVRSIVATQGTNEWIKLVS 657
QY 411 410
Db 658 KLLSQIGIPHQTYQYQENGKDRGYILEITGKDLILFQTLGIFGERKNALLKAIS 717
QY 411 410
Db 718 QREMANLENGYRSEFNVSVEYEGVKYDUTLEGTPYFANGILTTHNSLYPSIIITHN 777
QY 421 VSPDTLEREGCKNDVAPVIGYKFKDFGFIPIISILGELITMRQIKKKRATIDPIBK 480
Db 778 VSPDTLNRGCKEYDVAPOVGHFCKDFGFIPIISILGELITMRQIKKKRATIDPIBK 837
QY 481 MLDYRQAVKLHA 493
Db 838 LLDYRQAVKLHA 897
QY 494 493
Db 898 VSGLEVPSENNRTNKAELKRVKALIRHDYSGKVTIRLKSRRITKSHSLFSVRNGEL 957
QY 494 493
Db 958 VEVTDGLKPGDLVAVPRELPEPNNHVLNVLVLLGTPEETLIVMTIPVKGKKNFFK 1017
QY 494 493
Db 1018 GMLRTLRIWIFGEKRPRTARRVIRHLEDLYVRLKIGYEVLDWLSKNRYRLYZALVEN 1077
QY 494 493
Db 1078 VRYNGNKBVYLFVENSIRDAVGIMPLKELKWKIGTLNGFRMRKLIEVDES LAKLLGYV 1137
QY 494 493
Db 1138 SEGAYARKQNPKNKWSYVKLYNEDPEVLDDEMLASRPFQKVRGRNRYVEIPKIGVLL 1197
QY 494 493
Db 1198 FENMCGVLAENKRIPEFVFTSPKGVRLAFLEGYFGDGDVHNPKRILSLTKSELLANQLV 1257

QY 494 493
Db 1258 LLLNSGVSAVXLGHDGSGVYRVYINBELPFVKLDKKKNAYSHVIPKVLSEVFGKVFQK 1317
QY 494 493
Db 1318 NVSPQIFRKWVEDGRDLDPKQAQRSLWLBEGDVVLDRVESVDYDGYVYDLSVEDNEF 1377
QY 494 493
Db 1378 LVGFGLYAHNSYGYGYARAEWYCKEASVTAWGREYITMTIKIEIEBKGYFKVIYS 1437
QY 544 TDGFYATIPCEKPEKTIKKAKELKYNISKLPGLLELEYEGFYLRGFFVAKKAVYIDEE 603
Db 1438 TDGFFATIPADAEYTKKAWELKYNISKLPGLLELEYEGFYLRGFFVAKKAVYIDEE 1497
QY 604 GRITTRGLEVRDMSIEIAKETQAVLEALLKQDGVKAVRIIVKEVTEKLSKYEVPEKL 663
Db 1498 GKITTRGLEVRDMSIEIAKETQAVLEALLKQDGVKAVRIIVKEVTEKLSKYEVPEKL 1557
QY 664 VIHEQTKDLSSEYKAIGPHVAIAKLAAGIKVPGCTIISYIVLRGSGKISDRVILLSEY 723
Db 1558 VIHEQTRDKYKATGPHVAVAKLAAGVVKIRPGTVISYIVLRGSGRIGRAIFDEF 1617
QY 724 DPXKHKYDDPYIENQVLPVLRILEAFGYRKEDLYQSSKQVGLDAWLK 773
Db 1618 DPTKHKYDAEYIENQVLPVLRILEAFGYRKEDLYQSSKQVGLDAWLK 1667
RESULT 11
DPOL THEG8 STANDARD; PRT; 1699 AA.
AC Q9HH84;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) (Contains: Endonuclease PI-Tspge8I
DE (EC 3.1.1.-) (Tsp-GE8 pol-1 intein); Endonuclease PI-Tspge8II
DE (EC 3.1.1.-) (Tsp-GE8 pol-2 intein)).
GN POL OR POL-1.
OS Thermococcus sp. (strain GE8).
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=105593;
RN [1]
RP SEQUENCE FROM N.A.
RA Querrellou J.J.E., Cambon M.A., Lesongeur F., Barbier G.;
RT "Thermococcales taxonomy and phylogeny based on the comparative use of
RT 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
RT genes";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3: TO 5' EXONUCLEASE ACTIVITY (BY SIMILARITY).
CC -!- FUNCTION: PI-TSPG8I AND PI-TSPG8II ARE ENDONUCLEASES
CC (POTENTIAL).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}[N].
CC -!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTENIN)
CC FOLLOWED BY PEPTIDE LIGATION.
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -!- SIMILARITY: IN THE INTENIN SECTION; BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

```

DR EMBL; AJ250333; CAC12850.1; -.
DR HSSP; P56689; 1TGO.
DR InterPro; IPR006172; DNA_pol_B.
DR InterPro; IPR006134; DNA_pol_B_dom.
DR InterPro; IPR006133; DNA_pol_B_exo.
DR InterPro; IPR003586; Hedgehog_HintC.
DR InterPro; IPR003587; Hedgehog_HintN.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR006142; INTEIN.
DR InterPro; IPR004042; Intein_endonuc.
DR InterPro; IPR004578; Pol2.
DR Pfam; P00136; DNA_pol_B_3.
DR Pfam; P03104; DNA_pol_B_exo; 1.
DR PRINTS; PR00379; INTEIN_N_TER; 2.
DR SMART; SM00305; HintC; 2.
DR SMART; SM00486; POLB; 1.
DR TIGRFAMs; TIGR01443; intein_Cterm; 2.
DR TIGRFAMs; TIGR01445; intein_Nterm; 2.
DR TIGRFAMs; TIGR00592; pol2; 2.
DR PROSITE; PS00116; DNA_POLYMERASE_B; FALSE_NEG.
DR PROSITE; PS0818; INTNIN_C_TER; 2.
DR PROSITE; PS0819; INTNIN_ENDONUCLEASE; 2.
DR PROSITE; PS0817; INTNIN_N_TER; 2.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Hydrolyase; Nuclease; Exonuclease; Multifunctional enzyme;
KW Protein splicing; Autocatalytic cleavage; Endonuclease; Intron homing.
FT CHAIN 1 491 DNA_POLYMERASE_1ST PART.
FT CHAIN 492 1026 ENDONUCLEASE_PI-TSPG881.
FT CHAIN 1027 1075 DNA_POLYMERASE_2ND PART.
FT CHAIN 1076 1464 ENDONUCLEASE_PI-TSPG881I.
FT CHAIN 1465 1699 DNA_POLYMERASE_3RD PART.
SQ SEQUENCE 1699 AA; 197323 MW; F38994351F0B12D3 CRC64;

Query Match 68.8%; Score 2769.5; DB 1; Length 1699;
Best Local Similarity 35.5%; Pred. No. 6.5e-143;
Matches 602; Conservative 85; Mismatches 83; Indels 927; Gaps 4;

QY 1 MIFDTYITKDGKFIIRIFKENGFEKLEDPHOPFYIYALLKODSAIDETKATKGERHG 60
DB 1 MILDTYITTEDGKPIRVKENGGEFKIEYDRNFEFYIYALLKODSAIEEVKKTAKRHG 60

QY 61 KIVRVVDAVKKKFLGRDVEVVKLIIFEHPQVPALRGKIREHPAVIDIYEYDIPFAKRY 120
DB 61 TVVVKVRAEKVKKKFLGRPIEVVKLYFTHPQVPALRGKIREHPAVIDIYEYDIPFAKRY 120

QY 121 LIDKGLIPVEGDEELKMAFDIETVHEGDEFGKEIIMI SYADEEAEARVITWKNIDLPY 180
DB 121 LIDKGLIPNEGDEKLMALFDIETLYHEGEEFAEGPIIMI SYADEEGARVITWKKVDUPY 180

QY 181 VDVSNEREMIKRFVQIVREKDPDLVITYNGDNFOLPYLIKRAEKLGVTLGLGRDKHPE 240
DB 181 VDVSSTKEMIKRFVLRVWEKDPDLVITYNGDNFDPAYLKRSEKLGKVFILGRDGS--E 238

QY 241 PKTHRMGDSFAVEIKGRIHFDLPVVRRTINPTTLEAVYEAIVLGKTSKLGAEETAAI 300
DB 239 PKIQRMGDRFAVEIKGRIHFDLPVVRRTINPTTLEAVYEAIVLGKTSKLGAEETATA 298

QY 301 WETERSKKLAQVSMEDARATVELGKEFPFMAELAKIGOSVMDVSRSSGTGNLVWYLL 360
DB 299 WETGEGLERVARYSMEDAKVTVELGKEFPFMAELAKIGOSLMDVSRSSGTGNLVWFL 358

QY 361 RVAYERNELAPNKPDEEVRRLRTTYLGGYVKEPERGLWENITVDFRCYLPSTIITHN 420
DB 359 RYAYERNELAPNKPDERELARR-RQSYAGGYVKEPERGLWNNIVLDFRSYLPSTIITHN 417

QY 421 VSPDTTLRECKNVDVAPVIGVYKCFKDFGFGFTIPSLGELITWQSIKKKMKATIDPEKK 480
DB 418 VSPDTTLRECKEYDVAPOVGHKCFKDFGFGFTIPSLGDLLEERQIKRMRATIDPEKK 477

QY 481 MLDYRQRAVKLHA----- 493
DB 478 LLDYRQRAIKILANSILPDENPLLVNGRLKLVIRIGDFVDNTWKKQPLENDGTVELEVS 537

```

```

QY 697 REGTISIVLGGSKISDRVLLSEYDPKXKHYDPDYIYENQVLPVAVLRILEAFGYRKE 756
Db 1618 REGTVISVLGGSGRIGRAIPFDFEFAKHYDAEYIYENQVLPVAVLRILAFGYRKE 1677
QY 757 DLKYQSSKQVGLDAMLK 773
Db 1678 DLRYQKTKVGLGAMLK 1694

RESULT 12
DPOL THEHY
ID QPOL THEHY STANDARD; PRT; 1668 AA.
AC Q9HH05;
DT 16-OCT-2001 (rel. 40, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) [Contains: Endonuclease PI-ThyII
DE (EC 3.1.-.-) (Thy pol-1 intein); Endonuclease PI-ThyI (EC 3.1.-.-)
DE (Thy pol-2 intein)] (fragment).
GN POL.
OS Thermococcus hydrothermalis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OC NCBI_TaxID=465339;
RN [1]
RP SEQUENCE FROM N.A.
RA Querellou J.J.E., Cambon M.A., Lesongeur F.O., Barbier G.;
RT "Thermococcales taxonomy and phylogeny based on the comparative use of
RT 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
RT genes.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION OF PI-THYI.
RX MEDLINE=20512590; PubMed=11058140;
RA Saves I., Bleau H., Dietrich J., Masson J.-M.;
RT "The Thy pol-2 intein of Thermococcus hydrothermalis is an
RT isochizomer of Pi-thyI and Pi-Tfulli endonucleases.";
RL Nucleic Acids Res. 28:4391-4396(2000).
CC -!- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY (BY SIMILARITY).
CC -!- FUNCTION: PI-THYI AND PI-THYII ARE ENDONUCLEASES. PI-THYI CLEAVES
CC THE INTEINLESS SEQUENCE OF THE THY DNA POL GENE. IT REQUIRES A 21-
CC BP MINIMAL RECOGNITION SEQUENCE.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -!- [PM]: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION (potential).
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -!- SIMILARITY: IN THE INTEIN SECTION, BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ245619; CAC18555.1; ..
CC HSSP; P56689; 1TGO.
CC REBASE; 4832; PI-ThyI.
CC InterPro; IPR006172; DNA pol B.
CC InterPro; IPR006134; DNA pol_B_dom.
CC InterPro; IPR006133; DNA pol_B_exo.
CC InterPro; IPR003866; Hedgehog_hintc.
CC InterPro; IPR003587; Hedgehog_hintc.
CC InterPro; IPR006141; Intein.
CC InterPro; IPR006142; Intein.
CC InterPro; IPR004042; Intein_endonuc.
CC InterPro; IPR004578; Pol2.
CC Pfam; PF00136; DNA_pol_B; 3.

```

```

DR Pfam; PF03104; DNA_pol_B_exo; 1.
DR PRINTS; PRO0379; INTEIN.
DR SMART; SM00305; HintC; 2.
DR SMART; SM00306; HintN; 2.
DR SMART; SM00486; POLBc; 1.
DR TIGRfams; TIGR01443; intein_Cterm; 2.
DR TIGRfams; TIGR01445; intein_Nterm; 2.
DR TIGRfams; TIGR00592; pol2; 2.
DR PROSITE; PS00116; DNA POLYMERASE_B; FALSE_NEG.
DR PROSITE; PS00818; INTEIN_C_TER; 2.
DR PROSITE; PS00819; INTEIN_ENDONUCLEASE; 2.
DR PROSITE; PS00817; INTEIN_N_TER; 2.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Hydrolase; Nuclease; Exonuclease; Multifunctional enzyme;
KW Protein splicing; Autocatalytic cleavage; Endonuclease; Intrin homing.
FT CHAIN 1 1
FT CHAIN 1 458 DNA POLYMERASE, 1ST PART.
FT CHAIN 459 995 ENDONUCLEASE PI-THYI (POTENTIAL).
FT CHAIN 996 1044 DNA POLYMERASE, 2ND PART.
FT CHAIN 1045 1433 ENDONUCLEASE PI-THYII (POTENTIAL).
FT CHAIN 1434 1668 DNA POLYMERASE, 3RD PART.
SQ SEQUENCE 1668 AA; 193319 MW; 5EEB805FEEDA71C8 CRC64;

Query Match 63.0%; Score 2535.5; DB 1; Length 1668;
Best Local Similarity 33.6%; Pred. NO. 3.3e-130;
Matches 560; Conservative 85; Mismatches 92; Indels 929; Gaps 4;

QY 34 FQYIVALLKDDSAIDKAIKGERHGKIVRVVDVAVKKVKKLGRDVEYWKLIFFHPQDV 93
Db 1 FEYIVALLKDDSAIDKAIKGERHGKIVRVVDVAVKKVKKLGRDVEYWKLIFFHPQDV 60
QY 94 PALRGKIREHPAVIDIYEDIPFAKYLIDKGLIPMEGDEELKMAFDITFTYHEGDEFG 153
Db 61 PAIRDEIRRHSAVVDIYEDIPFAKYLIDKGLIPMEGDEELKMAFDITFTYHEGDEFG 120
QY 154 KGEIIMISYADEEAEARVITWKIDIPYVDVNEREMIKSFQIVREKDPDVLITNGDN 213
Db 121 TGPILMISYADEEAEARVITWKIDIPYVDVNEREMIKSFQIVREKDPDVLITNGDN 180
QY 214 FDLFYLIKRAEKLGVTLLGRDKEHPKIHMGSGSFAYEIKGRHFDLPFVVVRTINLP 273
Db 181 FDPAYLKKEKIGIKFTLRDGS--EPKIQMGDRFAVEVKGRIHFDLPVIRRTINLP 238
QY 274 TYTLEAVYEAIVGKTKSLGAEIIAIWETESMKLAQYSMEDARATVELGKEPPEMEA 333
Db 239 TYTLEAVYEAIVGKTKSLGAEIIAIWETESMKLAQYSMEDARATVELGKEPPEMEA 298
QY 334 ELAKLIGQSVWDVSRSTGNLVEWYLLRVAYERNELAPNKPDEEYRRRLRTTYLGGYVK 393
Db 299 QLSRLIGQSLWDVSRSTGNLVEWYLLRVAYERNELAPNKPDEEYRRRLRTTYLGGYVK 357
QY 394 EPERGLWENITYLDFPCLPSYIIVTHNVSPDTLREGCKQYDVAVIVGKFKDPPGFIP 453
Db 358 EPERGLWENITYLDFPCLPSYIIVTHNVSPDTLREGCKQYDVAVIVGKFKDPPGFIP 417
QY 454 SILGELITVRQETIKKKWKATIDPIEKMKLDYQRAVKLHA----- 493
Db 418 SLGALDERQKIKKKWKASIDPLEKLDYQRAVKLHA----- 477
QY 494 ----- 493
Db 478 RICEFVDFKLMETDSELVKNGDTEVLEVRGIRALSPDRKSKARVMFKAIRHRYSGDV 537
QY 494 ----- 493
Db 538 YEIVLGSRRITVTEGHSIFAYGDGELREVTGGEIKAGDLLAVPRVNLPEKKERLNIVE 597
QY 494 ----- 493
Db 598 LLRLPEETGDIIITIPVKGKKNPFKMLRTLRLWISGEKRPRTARRYLEHLEGLGYVR 657
QY 494 ----- 493

```



```

Db 658 LKKIGYVTDREGLRYKLYERLVEAVRYNGKREYLVEFNAVRDVIALMPEELRDWL 717
QY 494 ----- 493
Db 718 VGTNGFRMPFVEIBEDFAKLLGYVSEGNARKWENQNGNSYTVKLYNQRVLDDME 777
QY 494 ----- 493
Db 778 SLAERFFRVKRGKNIETPRKWAYIIFENLCOTLAENKRVPEAIFTSPEVRFVAFIEG 837
QY 494 ----- 493
Db 838 FIGDGVHPSKRVRLSTKSELLVNLGLVLLNSLGVSAIKIRHDSGVYRVYVNEELPPTDY 897
QY 494 ----- 493
Db 898 RKKQNAVYSHVIPKEILEETFGKVFORSYERFRELKSEKLDGKAKRIEVLNGDVV 957
QY 494 ----- 516
Db 958 LDKVLEKKRPYEGYVYDLSVEEDENFLAGFLLYAHNSYGYGYARARWYCKECAESV 1017
QY 517 TAMGRHYIEMTIKIEBKFGKVLAD----- 543
Db 1018 TAMGRDYIETTHIEERFGKVLADSVTGETETETETETETETETETETETETETET 1077
QY 544 ----- 543
Db 1078 EKECVLEGVEALTLDNRGLVWKSVPYVMRHTNKRIYRWNTSNWYLDVTDHSLIGY 1137
QY 544 ----- 543
Db 1138 MNTSKVPGKPKERLVEVKPGELGESVKSLLITPNRAIAHGIRVNPVIAVKLWELLGLLVG 1197
QY 544 ----- 543
Db 1198 DGNWGGOSNAKYNVGLSLDKBEEIKILPKLNTGIISNYDKSKGDVLSILSKWLA 1257
QY 544 ----- 543
Db 1258 RPYRYFKDESGSKRIPEFMFNLPREYIEAFLRGLFSADGTVSLRGVPEVRLTSVNP 1317
QY 544 ----- 543
Db 1318 SSSVRKLLMWGVNSMFMETPNPNRYLGKESGTHSVHVRKDKHRAERIGFLLDRKATK 1377
QY 544 ----- 547
Db 1378 LSENLGHTSKKRAKYDFDLVYPKVEEIAVDGVVYDIEVEGTHRFANGILVHNTDGF 1437
QY 548 YATIPGEKPEPTIKKAKKFLKINSKLPGLLELEYEGFYLRGFPVAKRYAVIDEGRIT 607
Db 1438 FATIPGADAETVKKKAKKFLKYNKAKLPGLLELEYEGFYVRGFFVTKKXAVIDEGRIT 1497
QY 608 TGLGVWRDSEIAKETQAKVLEAILKEDSVKAVEIVKDWBEIAKYQVPLEKVIHE 667
Db 1498 TGLGIVRDSEIAKETQARVLEAILRHGDVEAVR-VKDVTEKLSKYVPEPKVIHE 1557
QY 668 QITKDLSEYKAGPHVAIAKLAAGIKVRGTTIISYIVLRGSKSIDRVILLSEYDPKK 727
Db 1558 QITRELKDYKATGPHVAIAKLAARGIKIREGTVISYIVLRGSGRIGDRAIPFDFPTK 1617
QY 728 HKYDPDYIENQVLPAVRLIIEAFGRKEDLKYSQSKQVGLDAMLK 773
Db 1618 HRYDAEYIENQVLPAVERILKAFGYKKEELRYQKTRQVGLGAWLK 1663

RESULT 13
DPOL.METVO
ID DPOL.METVO STANDARD; PRT; 824 AA.
AC P52025;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)

```

```

28-FEB-2003 (Rel. 41, Last annotation update)
DNA polymerase (EC 2.7.7.7).
POL.
Methanococcus voltae.
Archaea; Euryarchaeota; Methanococci; Methanococcales;
Methanococcaceae; Methanococcus.
NCBI_TaxID=2188;
[1]
SEQUENCE FROM N.A.
MEDLINE=95014087; PubMed=7929013;
Konisky J., Paule S.M., Carinato M.B., Kansy J.W.;
"The DNA polymerase gene from the methanogenic archaeon Methanococcus
voltae.";
J. Bacteriol. 176:6402-6403 (1994).
-|- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ [DNA] (N).
-|- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; L33366; AAA72443.1; -.
HSP; P56689; 1TGO.
InterPro; IPR006172; DNA_pol_B.
InterPro; IPR006134; DNA_pol_B_dom.
InterPro; IPR006133; DNA_pol_B_exo.
InterPro; IPR004578; Pol2.
Pfam; PF00136; DNA_pol_B; 1.
Pfam; PF03104; DNA_pol_B_exo; 1.
PRINTS; PR0106; DNAPOLE.
SMART; SM00486; POLB; 1.
TIGRFAMs; TIGR00592; pol2; 1.
PROSITE; PS00116; DNA_POLYMERASE_B; 1.
Transferase; DNA-directed DNA polymerase; DNA replication;
DNA-binding.
KW SEQUENCE 824 AA; 96754 MW; 094A630D8C1514FC CRC64;
Query Match 34.6%; Score 1392; DB 1; Length 824;
Best Local Similarity 36.1%; Pred. No. 1.9e-68;
Matches 308; Conservative 179; Mismatches 251; Indels 116; Gaps 21;
QY 4 DTDYITKQKPIRIIFKENG-----EFKIELDPHQPIYIALLKDDSAI-----D 49
Db 2 DLDYNSKD--LCIDMYKNCGLKKEPEINLQKECEKPYFYVDTSEPKREIYDYLGLNOEI 59
QY 50 EIKATKGERHG---KIVRVVDVAVKVKKFL-----GRDV-----EYWKLI 86
Db 60 DLKLEPEPEPNTSLKVQDLITNIEIKIYVSDVILNGKDISEVDFDKKKERKICKVY 119
QY 87 FEHPQDVPALRKIEHFAVIDIYDYIPPAKRYLIDKGLIPM-----EGDEELK 136
Db 120 VKYPNHVKIIRYFKEFG---KSYEFDPFLRRYIMIDQDIVPSAKYSEDNKIDNSIPELN 176
QY 137 LMAPILETYHSGDEFGKEGELIMISYADBEERAVITKNIDLP-----YDVVVSNEREMIK 192
Db 177 CLAFDMELYCKEPEAKKADPIIMVNLFSQDYKVITYKFFENSEYNGCYDYVYKDEKELIQ 236
QY 193 RFVQIVREKDPVLITYNGDNFDLPYLIRAKELGVTLILGRDKHEPPE---KIRMGDS 249
Db 237 KTEIL--KQYDVIVTYNGDNFDPYLLKCRANIVIELDFDNASNSQQPQIIKISKGIN 294
QY 250 FAVEIKGRITHFDLPVVRRTINLPTVLEAVVGLTKSKLGAEEIAAIWETESMEKK 309
Db 295 RKSKIPGIITHIDLYIARLKLNLTKYLENVVQELFKINKEAVDYGDPKMWETEDT--T 352
QY 310 LAQYSMEDARATYELGKEFFPMEEALAKLIGOSVMDVSRSSTGNLVYVLLVAVERNEL 369
Db 353 LLRYAYEDALYTYMGNYVFLPLEIMPSRVNQNPFLYDTSRMNSQWVEFILLKRSPEQNMI 412

```

```

QY 370 APNKPDEEYRRRLRTTYLGGVYKPEPGLWENITLDFRCLYPSIIIVTHNVSPDTL--E 427
Db 371 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 413 SNRPSSSSYRRAKFSYEGGVREPLKGIQEDIVSLDFMSLYPSILSHNISPETVIYE 472
QY 428 REGKNYDVAPIGVYKFKDFPGFIPSPILGELITWRQEIKKOKKATIDPISK-----K 480
Db 473 EKERENNEL-----GIIPKTLNELLSSRRKHIMLLK--DKIQNEFDEEYS 516
QY 481 MLDYRQRAVKLHANSYGYMGYPKARWYKCAESVTAWRGHHYIMTIKEIEEKFGFKVL 540
Db 517 RLEHRSQKSLVLANSHYGLAFPMARWYSDKAEMVTLGRKYIQTETIEKAE--FGFKVI 575
QY 541 YATDGFVATIPCEPETIKKK-----AKFELKYNSKLPGLL 578
Db 576 YATDGFVAKWYDYLQKQKEENDSKSLNPKLSKEELIITLTKFKLGINBELPGM 635
QY 579 ELEVEGFYLRGPFVAKRYAVIDEGRITTRGLEVVRRDWSIEAKETQAKVLEAILKEDS 638
Db 636 ELEFGHFKRGLFVTKKVALIEDDGHIVKGLEVVRRDWSNIAKDTQOAVIRALLEDDG 695
QY 639 VEKAVEIVKDVVEEIAKYQVPLEKLVIIHQITKDLSEYKAIGPHVAIAKRLAAGKIVRP 698
Db 696 VNLAKKIINTDNLKKNIDKNDLLINTQLTNTEYKSTAPHIEVAKIKQKQDSVRV 755
QY 699 GTIISVILRGSGKSLDRVILLSEYDPKKGKYPDPYIENOVLPVAILLEAFGRKEDL 758
Db 756 GDVISYIIVKGRSISERAELL-EY---AGDYDINIYIDNQLVLPVIRIMESLSGISEDEL 811
QY 759 KYQSKQKQVLDAWL 772
Db 812 K-NSGQFKLDQFM 824

RESULT 14
DPOL ARCFU
ID DPOL ARCFU STANDARD; PRT; 781 AA.
AC 029753;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (SC 2.7.7.7).
GN POL OR POLB OR AFO497.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OX Archaeoglobaceae; Archaeoglobus.
RN [1] TaxID=2234;
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370 (1997).
CC -! CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -! SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AEO01070; AAB90741.1; -
CC PIR; A69312; A69312.
CC HSP; P56689; ITGO.
CC TIGR; AFO497; -
CC InterPro; IPR0061172; DNA_pol_B.
CC InterPro; IPR0061134; DNA_pol_B_dom.
CC InterPro; IPR0061133; DNA_pol_B_exo.
CC InterPro; IPR004578; Pol2.
CC Pfam; PF001136; DNA_pol_B; 1.
CC Pfam; PF031104; DNA_pol_B_exo; 1.
CC SMART; SM00486; POLBc; 1.
CC TIGRFAMs; TIGR00592; pol2; 1.
CC PROSITE; PS00116; DNA_POLYMERASE_B; 1.
CC Transferrase; DNA-directed DNA polymerase; DNA replication;
CC DNA-binding; Complete proteome.
CC SQ SEQUENCE 781 AA; 89850 MW; 761C7B18FCC61B8B CRC64;

Query Match 32.3%; Score 1301; DB 1; Length 781;
Best Local Similarity 37.4%; Pred. No. 1.6e-63;
Matches 305; Conservative 155; Mismatches 266; Indels 90; Gaps 21;

QY 2 IFPTDYITKDGKPIIRIFKKEGFEKIELDPHPQPIYIALLKDDSAIDRIKAKGERHGK 61
Db 8 LIDADYETIGGKAVVRLCKDQGIQFVAYDYNDFPYFYVIGVDE--DILKNAATSTRE 64
QY 62 IVRVDAVVKVKKFLGRDVEVMKLIPEHPQDYPALRGKIREHPAVIDIEYDIPFAKRYL 121
Db 65 VIKLSEFKAQLTKLGREVEGYIVVAHHPQHPKURDYLSQFG--DVEADIPFAIRYL 121
QY 122 IDKGL-----IPMEGDE-----ELKMAPDIEFTYHEG-DEFG 153
Db 122 IDKDLACMDGIAEGEKQGVIRSYKIEKVERIPMERPELAKMLVDCMLSSFGMPPE 181
QY 154 KGEIIMISYADBEARVITWKIDLDYVDVNEREMIKRFQIVREKDPVLITNGN 213
Db 182 KDPFIIVISVKTNDDEIIL-----LTGDERKIIISDFVKLIKSYDPDIIVGNQDA 230
QY 214 FDLPLYIKRAEKLGVTLILGRDKHEPPEPKIHMWGSPFAVEIKGRHFDLPFVVRRINLP 273
Db 231 FWPYLRKAEARNWNIPLDVGRDGSN---VVRGG---RPKITGRNLVDIYDIAMRISDK 284
QY 274 TYTLEAVYEAVLGKTK-SKLGAEETAAIWETESMKKLAQYSMEDARATYELKGKFFPME 332
Db 285 IKKLENVABFLGTGKIBIADIEAKDIYRWSRGEK-EKVLNVARQDAINTYIAKELLPMH 343
QY 333 AELAKLIGQSVMDVSRSSSTGNLVVYLLRVAVERNELAPNKPDEBEYRRRLRTTYLGGV 392
Db 344 YELSKMIRLPVDDVTMGKQKQVDMLLSEAKKIGELIAPNPEHAE-----SYEGAFV 396
QY 393 KPEPGLWENITYLDPRCLYPSIIIVTHNVSPDLRECGKN--YDVAPIVGYKFKDPPG 450
Db 397 LEPEPGLHENVACLOFASPMYSIMIAFNISPDY---GCRDDCYE-APEVGHKFRKSPDG 452
QY 451 FIPSLIGELITWRQELKKWK-ATIDPLEKMLDYRQRAVKLHANSYGYMGYPKARWYS 509
Db 453 FKRILRMLEIKRELKVELKNLSESEYKLLDIKQTLKVLITNSFYFGWGNLAWTC 512
QY 510 KCAESVTAWRGHHYIEMTIKEIEEKFGFKVLVADTGFYATIPGKPEPIKKAKEFLKY 569
Db 513 HPCAEATTANGRHFRITSK-IAESMGFKVLYGDTDSIFVTKAGMTKEDVDVDR----LIDK 567
QY 570 INSKLPGLLEVEGYLGFVAKRYAVIDEGRITTRGLEVVRRDWSIEAKETQAKV 629
Db 568 LHEELP--IQIEVEYISAIFVEKKRYAGLFGDLGRVVKGLEVRGDMCELAQKQREV 625
QY 630 LEAILKEDSVKAVITKDWVEEIAKYQVPLEKLVIIHQITKDLSEYKAIGPHVAIAKRL 689
Db 626 IEVILKEKNPEKALSLVKDVIIRIKEGKVSLEEVIVKGLTKPKSKYESMQHVAALKA 685

```

QY 690 AAKGKIVRPGTIIISYIVLRSGKISDR---VILLSEYDPKK-----HKYDPDY 735
 Db 686 REMGIIIVPSSKIGYIVKSGNIGRAVPIDLIEDFNGENRIKTKSGIEIKLKDXY 745
 QY 736 IENQVLPAVRIILAEGRKEDLYOSSQVGLDAW 771
 Db 746 IDNQIIPVLRILERFGYTEASLK--GSSQMSLDSF 779

RESULT 15
 DPOL METJA
 ID DPOL METJA STANDARD; PRT; 1634 AA.
 AC Q58295; 1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) [Contains: Mja pol-1 intein; Mja pol-2
 DE intein]
 GN POL OR MJ0885.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337993; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Corton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii";
 RL Science 273:1058-1073(1996).
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA} (N).
 CC -!- PGM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U67532; AAB98889.1; -.
 CC DR HSSP; P56689; 1TGO.
 CC DR TIGR; MJ0885; -
 CC DR InterPro; IPR006172; DNA_pol_B.
 CC DR InterPro; IPR006134; DNA_pol_B_dom.
 CC DR InterPro; IPR006133; DNA_pol_B_exo.
 CC DR InterPro; IPR003586; Hedgehog_hintC.
 CC DR InterPro; IPR003587; Hedgehog_hintN.
 CC DR InterPro; IPR006141; InteIn.
 CC DR InterPro; IPR006142; INTEIN.
 CC DR InterPro; IPR004042; InteIn_endonuc.
 CC DR InterPro; IPR004578; POL2.
 CC DR Pfam; PF00136; DNA_pol_B; 3.
 CC DR Pfam; PF03104; DNA_pol_B_exo; 1.
 CC DR PRINTS; PR00379; INTEIN.
 CC DR SMART; SM00305; HintC; 2.
 CC DR SMART; SM00306; HintN; 2.
 CC DR SMART; SM00466; POLBc; 1.
 CC DR TIGRFAMs; TIGR01443; intein_Cterm; 2.

DR TIGRFAMs; TIGR01445; intein_Nterm; 2.
 DR TIGRFAMs; TIGR00592; pol2; 1.
 DR PROSITE; PS00116; DNA POLYMERASE_B; 1.
 DR PROSITE; PS00818; INTEIN_C_TER; 2.
 DR PROSITE; PS00819; INTEIN_ENDONUCLEASE; 2.
 DR PROSITE; PS00817; INTEIN_N_TER; 2.
 DR TRANSFERASE; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Autocatalytic cleavage; Protein splicing;
 KW Complete proteome.
 FT CHAIN 1 425 POL, 1ST PART (POTENTIAL).
 FT CHAIN 426 794 MJA POL-1 INTEIN (POTENTIAL).
 FT CHAIN 795 882 POL, 2ND PART (POTENTIAL).
 FT CHAIN 883 1358 MJA POL-2 INTEIN (POTENTIAL).
 FT CHAIN 1359 1634 POL, 3RD PART (POTENTIAL).
 SQ SEQUENCE 1634 AA; 191708-MW; 841FAFAB1F97DDD CRC64;
 Query Match 38.1%; Score 1292; DB 1; Length 1634;
 Best Local Similarity 22.7%; Pred. No. 1.3e-63;
 Matches 374; Conservatave 159; Mismatches 217; Indels 894; Gaps 23;
 QY 2 IFDIDYITKDGKPIIRIKKENGGEFKIELDPHFOPYIYALLK---DSDAIDEIK--AIK 55
 Db 13 LIDNTYKTIEDKAVIYLYLINS-----ILKRDFFKPYFVELHKEKVENEDIEKIEKELLK 68
 QY 56 GERHGKIVRVVDVAVKVKKK-FLGRDVEVWKLIFEHQDVPALRGKIREHPAVIDIYEVDI 114
 Db 69 ND---LLKPVENIEVWVKIILRKEKEVIKIIATHPKQVKPLR-KIKECEIVKEIYEHDI 123
 QY 115 PFAKYILDKGLIPM-----EGDE-----ELKLMADFIDTEVHEGDEFGKGE-IIMIS 161
 Db 124 PFAKYILIDNIIIPMTYWFENKPKVPSEIEPLKSAVAFDMEVNTRDIEPNRDPFILMAS 183
 QY 162 YADSEARVITWKNIDLPVDVNSNEREMIKRFVQIVREKDPDYLITYNGDNFDPYLIK 221
 Db 184 FWDENGKGVITVYEFNHNIEVWVKELIKKIIETLKEY--DVIYTYNGDNFPPYLKA 241
 QY 222 RAEKLGVTLLLRDKEHPEPKIHMGDSFAVEIKGRIFHDFLPVVRKTINLPTVTLAVY 281
 Db 242 RAKYIGIDINLGKGE--ELKIKGGMGEYRYPGRVHIDILYPSRLLKULKYTLDEVDV 299
 QY 282 EAVLGKTKSKLGABEIAAIWETESMKLAQYSEMEDARATYELGKEFFPMEAEAKLIGQ 341
 Db 300 YNLFGIEKLIKPHTKIVDYWANND--KTLIEYSIQDAKYTYKIGYFFELVEMFSRIWQ 357
 QY 342 SWDVSRSTGNLVEWYLLRVAYERNELAPNKPBEERYRRLRTTYLGGYKEBERGLWE 401
 Db 358 TPEFETRMSGQWVEYLLMKCAFKNMIVPNKPDDEEYRVRRLVTTTTEGGYKEPEKGMFE 417
 QY 402 NITYLDPRC----- 410
 Db 418 DIISMDFRCHPKGTWVVKGVNIEDVKEGNYVLGIDGKQVKVKWKVYEGEGLINVN 477
 QY 411 ----- 410
 Db 478 GLKCTPNHKIPURYKIKHKIKNDYLRDIYAKSLLPKSGKLLCKDPFETIGNYEK 537
 QY 411 ----- 410
 Db 538 YINDMDEDEILKSELIGILLAEGLLRDRDIEYFDSRCKKRLSHQRYVEITVNEDEKDFI 597
 QY 411 ----- 410
 Db 598 EKIKYIFKKLFNYLYVRKKGTKAITLGCAKQDIYLIKIEILKNKYEKLPNAILRGFFE 657
 QY 411 ----- 410
 Db 658 GDGYVNTVRVAVVNGTNNYDKIKFIASLLDRGLGIKYSFYTSYVEERGGKLLKRYVIEIF 717
 QY 411 ----- 410
 Db 718 SKGDLIKFSILISFISRRKNNLLNIEIROKLYKIGIDYFDLDDVDCVSLSEYKGEVYDL 777
 QY 411 -----LYPSIIIVHNVPDTLREGCKNYDVA-PIVGYKFKCKDFGCF 451